

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:40 ; Search time 12.371 Seconds
(without alignments)
1059.473 Million cell updates/sec

Title: US-09-135-238B-2_COPY_273_390

Perfect score: 616

Sequence: 1 KRAVERKALSRRLAVR.....HQPAAMMEDSDSDYINVPA 118

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	390	20	AA42225 Human Toso protein
2	616	100.0	390	20	AA17496 Human Toso protein
3	616	100.0	390	20	AA05001 Human FIGRL-1 prot
4	356.5	57.9	422	22	AAE05349 Mouse Toso protein
5	93	15.1	407	21	AAAG40735 Zea mays protein f
6	91.5	14.9	538	22	AAAB82806 Human low density
7	90	14.6	416	22	AAU17493 Novel signal trans
8	88	14.3	609	21	AAU53876 A human seven tran
9	87	14.1	314	22	AAU32408 Novel human secret
10	87	14.1	609	22	AAU08993 Human G protein-co
11	87	14.1	609	22	AAAG4292 Human GTP-binding

12	87	14.1	740	22	AAAG3401 Human polypeptide,
13	86	14.0	343	22	ABG12999 Novel human diagn
14	85.5	13.9	2205	16	AAAG79048 Infectious rubella
15	85	13.8	213	22	AAAG90922 C glutamic prote
16	85	13.8	554	22	ABB71946 Drosophila melanog
17	84.5	13.7	217	22	ABB10356 Human cDNA SEQ ID
18	84.5	13.7	358	21	AAU94336 Human cell surface
19	84.5	13.7	359	20	AAU41690 Human PRO329 prote
20	84.5	13.7	359	21	AAAB44246 Human PRO329 (UNQ2
21	84.5	13.7	359	21	AAAB34744 Human secreted pro
22	84.5	13.7	359	22	AAU29044 Human PRO polypept
23	84.5	13.7	359	22	AAE03451 Human gene 25 enco
24	84.5	13.7	550	22	AAAB82807 Rabbit low density
25	84	13.6	466	13	AAAB26505 Adrenalin receptor
26	84	13.6	609	22	AAE04548 Human G-protein co
27	83	13.5	706	22	ABBB66693 Drosophila melanog
28	83	13.5	1260	22	ABBB64653 Drosophila melanog
29	83	13.5	1740	21	AAU83017 Rat shank 3a. Rat
30	83	13.5	1740	22	AAAB31517 Amino acid sequenc
31	82.5	13.4	481	22	ABG05355 Novel human diagn
32	82	13.3	381	22	ABG89886 C glutamic prote
33	82	13.3	452	22	AAAB94445 Human protein sequ
34	82	13.3	681	22	ABG07912 Novel human diagn
35	82	13.3	802	22	ABB71892 Drosophila melanog
36	82	13.3	981	22	ABG22248 Novel human diagn
37	82	13.3	1115	22	ABG28242 Novel human diagn
38	81.5	13.2	2115	19	AAW59276 Rubella virus RA27
39	81	13.1	145	22	AAO06808 Human polypeptide
40	81	13.1	194	22	AAAG74948 Human colon cancer
41	81	13.1	292	22	ABG15356 Novel human diagn
42	81	13.1	385	21	AAAB56596 Human prostate can
43	81	13.1	550	16	AAAT2830 Human TIABP2. Hom
44	81	13.1	550	19	AAW64717 Human TIABP2 prote
45	81	13.1	550	20	AAU28589 TIA-1 binding prot

ALIGNMENTS

RESULT 1
AA42225
ID AA42225 standard; Protein; 390 AA.
XX
AC AA42225;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human Toso protein sequence.
XX
KW Human; Toso protein; target; drug screening; diagnosis; apoptosis;
XX
OS Homo sapiens.
XX
PN WO9950671-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06945.
XX
PR 30-MAR-1998; 98US-0050861.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Payan D;
XX
DR WPI; 1999-591379/50.
XX
N-PSDB; AA225422.
XX
PT Screening agents useful for modulating apoptosis and controlling
XX
apoptosis related diseases -
XX
Claim 1; Fig 2a; 75pp; English.

XX The present invention describes a method of Screening for a bioactive
 CC agent capable of binding a Toso protein. Also described a methods for:
 CC (1) screening a bioactive agent capable of modulating activity of a
 CC Toso cell-surface receptor, comprising adding a candidate bioactive
 CC agent to a cell comprising a recombinant Toso nucleic acid, and
 CC exposing the cells to an apoptotic agent that will induce apoptosis;
 CC (2) modulating apoptosis comprising administering an exogenous
 CC compound that binds Toso, to a cell; (3) identifying a cell containing
 CC a mutant Toso gene, comprising determining it's sequence; (4)
 CC identifying the Toso genotype, comprising determining the sequence of
 CC at least one Toso gene; and (5) diagnosing an apoptosis related
 CC condition, comprising measuring activity of Toso in a tissue, and
 CC comparing to the activity from non-affected individual's tissue, where
 CC a reduced activity of the patient indicates risk of an apoptosis related
 CC condition. The methods are useful for identifying agents capable of
 CC diagnosing and treating apoptosis related disease, their use for
 CC modulating apoptosis, and methods for diagnosing the disease state.
 CC The present sequence represents the human Toso protein for use in
 CC methods from the invention.
 XX
 SQ Sequence 390 AA;

Query Match 100.0%; Score 616; DB 20; Length 390;
 Best Local Similarity 100.0%; Pred. No. 3.le-53;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRAVERKALSRARRLAVRMRALESSQRPGRSPRPRSONNIYSACPRRARGADAAGTGE 60
 DB 273 KRAVERKALSRARRLAVRMRALESSQRPGRSPRPRSONNIYSACPRRARGADAAGTGE 332
 QY 61 APVPGCAPLPAPLQVSESPWLHAPSLKTSCEYVSLYHQPAAWMESSDDYINVPA 118
 DB 333 APVPGCAPLPAPLQVSESPWLHAPSLKTSCEYVSLYHQPAAWMESSDDYINVPA 390

RESULT 2
 AAY17496
 ID AAY17496 standard; Protein; 390 AA.
 XX
 AC AAY17496;
 XX
 DT 03-AUG-1999 (first entry)
 XX
 DE Human Toso protein.
 XX
 KW Toso protein; tumour necrosis factor mediated apoptosis inhibition;
 KW TNF mediated apoptosis; T cell overactivity; autoimmune disease;
 KW Sjogrens connective tissue disorder; transplant rejection; cancer.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..17
 FT /label= signal
 FT Protein 18..390
 FT /label= Toso
 FT Region 254..272
 FT /label= transmembrane_region
 XX
 PN WO9925832-A1.
 XX
 PD 27-MAY-1999.
 XX
 PF 16-NOV-1998; 98WO-US24391.
 XX
 PR 17-AUG-1998; 98US-0135238.
 PR 17-NOV-1997; 97US-0066063.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Hitoshi Y, Nolan GP;
 XX

DR WPI; 1999-338007/28.
 DR N-PSDB; AAX76123.
 XX
 PT DNA encoding Toso, a protein having inhibitory effects on TNF
 PT mediated apoptosis
 XX
 PS Claim 20; Fig 2a; 70pp; English.
 XX
 CC The present sequence is a Toso protein (I). (I) has anti-apoptotic
 CC and cytostatic activity. Toso (named after a Japanese liquor that is
 CC drunk on New Year's Day to celebrate long life and eternal youth) most
 CC likely acts by induction of cFLIP expression which inhibits caspase-8
 CC or to treat an apoptosis related condition in a mammal. Apoptosis
 CC related conditions can also be treated by administration of the Toso
 CC protein or antibody. Apoptosis related or mediated conditions that can
 CC be treated include diseases characterized by T cell overactivity, e.g.
 CC Sjogrens connective tissue disorder, autoimmune diseases, diseases where
 CC T cells actively destroy cells, including transplant rejection and
 CC conditions where cells of any kind that are not dying express Toso
 CC appropriately, e.g. cancer of T or B cell origin (where increased
 CC apoptosis would be appropriate).
 XX
 SQ Sequence 390 AA;

Query Match 100.0%; Score 616; DB 20; Length 390;
 Best Local Similarity 100.0%; Pred. No. 3.le-53;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRAVERKALSRARRLAVRMRALESSQRPGRSPRPRSONNIYSACPRRARGADAAGTGE 60
 DB 273 KRAVERKALSRARRLAVRMRALESSQRPGRSPRPRSONNIYSACPRRARGADAAGTGE 332
 QY 61 APVPGCAPLPAPLQVSESPWLHAPSLKTSCEYVSLYHQPAAWMESSDDYINVPA 118
 DB 333 APVPGCAPLPAPLQVSESPWLHAPSLKTSCEYVSLYHQPAAWMESSDDYINVPA 390

RESULT 3
 AAY05001
 ID AAY05001 standard; Protein; 390 AA.
 XX
 AC AAY05001;
 XX
 DT 16-JUN-1999 (first entry)
 XX
 DE Human PIGRL-1 protein sequence.
 XX
 KW PIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;
 KW X-linked Severe Combined Immunodeficiency; XSCID; Iga deficiency;
 KW diagnosis; therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP905238-A2.
 XX
 PD 31-MAR-1999.
 XX
 PF 14-AUG-1998; 98EP-0306487.
 XX
 PR 30-OCT-1997; 97US-0961564.
 PR 25-AUG-1997; 97US-0056935.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Sweet RW, Truneh A, Wu S;
 XX
 DR WPI; 1999-192666/17.
 DR N-PSDB; AAX28178.
 XX
 PT New polypeptides encoding PIGRL-1 useful for treating diseases such
 PT as X-linked Severe Combined Immunodeficiency
 XX

PS Claim 11; Page 7; 26pp; English.

XX This sequence is the human PIGRL-1 protein of the invention.

CC Autoimmune diseases involving altered expression or activity of PIGRL-1

PT may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined

CC Immunodeficiency (XSCID) and IGA deficiency. These diseases can be

CC diagnosed or susceptibility to them predicted by: (1) determining whether

CC there is a mutation in the genomic copy of the gene encoding PIGRL-1; or

CC (2) measuring the amount of PIGRL-1 in a sample derived from the patient.

CC Patients deficient in PIGRL-1 can be treated by administering either the

CC PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient.

CC Patients with excessive expression or activity of PIGRL-1 can be treated

CC by administering an antagonist of PIGRL-1, an antisense nucleic acid

CC molecule which inhibits the expression of PIGRL-1 or administering

CC sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can

CC be used to identify its agonists by contacting a cell expressing PIGRL-1

CC with a candidate compound in the presence of a signal system and noting

CC the candidate as an agonist if a signal is produced. The same method can

CC be used to identify antagonists of PIGRL-1 but the presence of an

CC antagonist is indicated by a decrease in production of the signal.

CC Antibodies against PIGRL-1 may be used to isolate or identify clones

CC expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to

CC identify chromosomal mutations in the gene encoding PIGRL-1 in patients.

CC This information may then be correlated with the incidence of autoimmune

CC disease in those patients to identify whether the mutation causes the

CC disease.

XX

SQ Sequence 390 AA;

Query Match 100.0%; Score 616; DB 20; Length 390;

Best Local Similarity 100.0%; Pred. No. 3.1e-53;

Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRAVERKALSRARRLAVRMRALESSQPRGSPRRSONNIYSACPRRARGADAAGTGE 60

Db 273 KRAVERKALSRARRLAVRMRALESSQPRGSPRRSONNIYSACPRRARGADAAGTGE 332

Qy 61 APVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHOPAAHMEDESDDDYNVPA 118

Db 333 APVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHOPAAHMEDESDDDYNVPA 390

RESULT 4

AAE05349

ID AAE05349 standard; Protein; 422 AA.

XX

AC AAE05349;

XX

DT 12-SEP-2001 (first entry)

XX

DE Mouse Toso protein.

XX

KW Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;

KW wound healing; immune response; vaccine; cancer; asthma; allergy;

KW cell trafficking; therapy; secreted protein; Fas-induced apoptosis;

Toso.

XX

OS Mus sp.

XX

PN WO200148192-A1.

XX

PD 05-JUL-2001.

XX

PF 21-DEC-2000; 2000WO-NZ00256.

XX

XX 23-DEC-1999; 99US-0171678.

PR 28-NOV-2000; 2000US-0724864.

XX

PA (GENE-) GENESIS RES & DEV CORP LTD.

XX

PI Watson JD, Murison JG;

XX

DR WPI; 2001-425665/45.

DR

DR N-PSDB; AAD10117.

XX Novel isolated polypeptide useful to isolate corresponding interacting

PT proteins or other compounds, to quantitatively determine levels of

PT interacting proteins or other compounds, and as therapeutic target -

XX

PS Claim 6; Page 78-79; 101pp; English.

XX

CC The patent discloses novel polynucleotides and their corresponding

CC proteins which play a major role in induction of growth, cell migration

CC and proliferation, cell-cell interaction and the differentiation of

CC tissue-specific cells. These proteins are important in the maintenance

CC of tissue integrity and thus are important in wound healing. They are

CC useful in various assays to determine the biological activity, to raise

CC antibodies, to isolate corresponding interacting proteins or other

CC compounds, to quantitatively determine levels of interacting proteins or

CC other compounds, and as therapeutic target in a whole range of disease

CC states. Compositions comprising the novel proteins of the invention are

CC useful for treating mammalian disorders. Polynucleotides of the invention

CC are useful in genome and physical mapping, in positional cloning of

CC genes, to tag or identify an organism or its reproductive material (as

CC non-disruptive tags for marking organisms), and for the diagnosis and

CC treatment of mammalian diseases which is the consequence of inappropriate

CC expression of kinase genes. They are useful for promoting immune response

CC as part of a vaccine or anti-cancer treatment, as target for cancer

CC treatment, as immunoregulatory and anti-inflammatory molecule, as

CC diagnostic for specific types of cancer and for development of an

CC anti-cancer treatment, and as a target for antagonists in the treatment

CC of diseases such as asthma and allergy. They are also useful to inhibit

CC or enhance the activity of the soluble molecule that binds proteins of

CC the invention, for tissue and neural regeneration, to promote or block

CC cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.

CC The present sequence is Toso, a secreted protein from mouse. Toso is

CC a cell surface, specific regulator of Fas-induced apoptosis in T-cells.

XX

SQ Sequence 422 AA;

Query Match 57.9%; Score 356.5; DB 22; Length 422;

Best Local Similarity 59.8%; Pred. No. 2.2e-27;

Matches 73; Conservative 12; Mismatches 32; Indels 5; Gaps 1;

Qy 1 KRAVERKALSRARRLAVRMRALESS-----QPRGSPRRSONNIYSACPRRARGADA 55

Db 284 KRAIQRRRRASSRRAGRLAMRRRGASRPPTQRRDAPQRPQRSONNIYSACPRRARGDPS 343

Qy 56 AGTGEAPVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHOPAAHMEDESDDDYN 115

Db 344 LGPAEAPLLNAPASASPQVLEAPWHPHPSLKMACEYVSLGYPVNLDPDSDDDYN 403

Qy 116 VP 117

Db 404 IP 405

RESULT 5

AAG40735

ID AAG40735 standard; Protein; 407 AA.

XX

AC AAG40735;

XX

DT 18-OCT-2000 (first entry)

XX

DE Zea mays protein fragment SEQ ID NO: 50584.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence; corn.

XX

OS Zea mays subsp. mays.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.
PF 99US-0121825.
XX 99US-0123180.
PR 99US-0123348.
PR 99US-0123548.
PR 99US-0125788.
PR 99US-0126264.
PR 99US-0126785.
PR 99US-0127462.
PR 99US-0128234.
PR 99US-0128714.
PR 99US-0129845.
PR 99US-0130077.
PR 99US-0130449.
PR 99US-0130510.
PR 99US-0130891.
PR 99US-0131449.
PR 99US-0132048.
PR 99US-0132407.
PR 99US-0132484.
PR 99US-0132485.
PR 99US-0132486.
PR 99US-0132487.
PR 99US-0132863.
PR 99US-0134256.
PR 99US-0134218.
PR 99US-0134219.
PR 99US-0134221.
PR 99US-0134370.
PR 99US-0134768.
PR 99US-0134941.
PR 99US-0135124.
PR 99US-0135353.
PR 99US-0135629.
PR 99US-0136021.
PR 99US-0136392.
PR 99US-0136782.
PR 99US-0137222.
PR 99US-0137528.
PR 99US-0137502.
PR 99US-0137724.
PR 99US-0138094.
PR 99US-0138540.
PR 99US-0138847.
PR 99US-0139119.
PR 99US-0139452.
PR 99US-0139453.
PR 99US-0139492.
PR 99US-0139454.
PR 99US-0139455.
PR 99US-0139456.
PR 99US-0139457.
PR 99US-0139458.
PR 99US-0139459.
PR 99US-0139460.
PR 99US-0139461.
PR 99US-0139462.
PR 99US-0139463.
PR 99US-0139750.
PR 99US-0139763.
PR 99US-0139817.
PR 99US-0139899.
PR 99US-0140353.
PR 99US-0140354.
PR 99US-0140695.
PR 99US-0140823.
PR 99US-0140991.
PR 99US-0141287.
PR 99US-0141842.
PR 99US-0142154.
PR 99US-0142055.
PR 99US-0142390.
PR 99US-0142390.
PR 08-JUL-1999;
PR 09-JUL-1999;
PR 12-JUL-1999;
PR 13-JUL-1999;
PR 14-JUL-1999;
PR 15-JUL-1999;
PR 16-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 19-JUL-1999;
PR 20-JUL-1999;
PR 20-JUL-1999;
PR 20-JUL-1999;
PR 21-JUL-1999;
PR 21-JUL-1999;
PR 21-JUL-1999;
PR 22-JUL-1999;
PR 22-JUL-1999;
PR 22-JUL-1999;
PR 22-JUL-1999;
PR 23-JUL-1999;
PR 23-JUL-1999;
PR 23-JUL-1999;
PR 26-JUL-1999;
PR 27-JUL-1999;
PR 27-JUL-1999;
PR 27-JUL-1999;
PR 28-JUL-1999;
PR 02-AUG-1999;
PR 02-AUG-1999;
PR 02-AUG-1999;
PR 03-AUG-1999;
PR 04-AUG-1999;
PR 04-AUG-1999;
PR 05-AUG-1999;
PR 05-AUG-1999;
PR 06-AUG-1999;
PR 06-AUG-1999;
PR 09-AUG-1999;
PR 09-AUG-1999;
PR 10-AUG-1999;
PR 11-AUG-1999;
PR 12-AUG-1999;
PR 13-AUG-1999;
PR 13-AUG-1999;
PR 16-AUG-1999;
PR 17-AUG-1999;
PR 18-AUG-1999;
PR 20-AUG-1999;
PR 20-AUG-1999;
PR 20-AUG-1999;
PR 23-AUG-1999;
PR 23-AUG-1999;
PR 25-AUG-1999;
PR 26-AUG-1999;
PR 27-AUG-1999;
PR 27-AUG-1999;
PR 27-AUG-1999;
PR 30-AUG-1999;
PR 31-AUG-1999;
PR 01-SEP-1999;
PR 07-SEP-1999;
PR 10-SEP-1999;
PR 13-SEP-1999;
PR 15-SEP-1999;
PR 16-SEP-1999;
PR 20-SEP-1999;
PR 22-SEP-1999;
PR 23-SEP-1999;
99US-0142803.
99US-0142920.
99US-0142977.
99US-0143542.
99US-0143624.
99US-0144005.
99US-0144085.
99US-0144086.
99US-0144325.
99US-0144331.
99US-0144332.
99US-0144333.
99US-0144334.
99US-0144335.
99US-0144352.
99US-0144632.
99US-0144884.
99US-0144814.
99US-0145086.
99US-0145088.
99US-0145085.
99US-0145087.
99US-0145089.
99US-0145192.
99US-0145145.
99US-0145218.
99US-0145224.
99US-0145276.
99US-0145913.
99US-0145918.
99US-0145919.
99US-0145951.
99US-0146386.
99US-0146388.
99US-0146389.
99US-0147038.
99US-0147204.
99US-0147302.
99US-0147192.
99US-0147260.
99US-0147303.
99US-0147416.
99US-0147493.
99US-0147935.
99US-0148171.
99US-0148319.
99US-0148341.
99US-0148565.
99US-0148684.
99US-0149368.
99US-0149175.
99US-0149426.
99US-0149722.
99US-0149723.
99US-0149929.
99US-0149902.
99US-0149930.
99US-0150566.
99US-0150884.
99US-0151065.
99US-0151066.
99US-0151080.
99US-0151303.
99US-0151438.
99US-0151930.
99US-0152363.
99US-0153070.
99US-0153758.
99US-0154018.
99US-0154039.
99US-0154779.
99US-0155139.
99US-0155486.

KW acquired immune deficiency syndrome.

XX OS Homo sapiens.
XX PN WO200154733-A1.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01312.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184684.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.


```
AAG64292
ID AAG64292 standard; Protein; 609 AA.
XX
AC AAG64292;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human GTP-binding protein-coupled receptor GPRv15.
XX
KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
KW muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;
KW G-protein.
XX
OS Homo sapiens.
XX
PN WO200148189-A1.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-JP09409.
XX
PR 28-DEC-1999; 99JP-0375152.
XX
PR 31-MAR-2000; 2000JP-0101339.
XX
PR 23-MAY-2000; 2000JP-0155978.
XX
PA (HELI-) HELIX RES INST.
XX
PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
PI Sugiyama T;
XX
DR WPI; 2001-425663/45.
DR N-PSDB; AAH49505.
XX
XX Family of guanosine triphosphate binding protein coupled receptors and
PT genes encoding them for treatment and prevention of diseases associated
PT with these receptors -
XX
PS Claim 1; Pages 68-72; 137pp; Japanese.
XX
CC The present sequence is the protein sequence for a human guanosine
CC triphosphate (GTP)-binding protein-coupled receptor. The receptor is
CC useful for the investigation, diagnosis, treatment and prevention of
CC diseases associated with GTP-binding protein-coupled receptors, including
CC neurological, circulatory, digestive system, immune system, muscle and
CC urinary system disorders. GTP-binding proteins are also known as
CC G-proteins.
XX
SQ Sequence 609 AA;
Query Match 14.1%; Score 87; DB 22; Length 609;
Best Local Similarity 29.6%; Pred. No. 2.2;
Matches 40; Conservative 8; Mismatches 31; Indels 56; Gaps 7;
QY 5 ERKAL-----SRARR-----LAVRRALESSQR-----PRGSPR-----PRS 38
|||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 434 ERRASLLAFADAPPSRARRRSAESLLSLRPSALDSGRGARDSPPGSPRRRRPGGPRS 493
QY 39 QN-----NIYSACPRRARGADAAGTGEAPVPGGAPLPAP 74
|||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 494 ASASLLPDFAFTAFCEFOALRRPPGPPFAAPADGDAD---PGEAPTTPSSAQRSGP 550
QY 75 LQVSESPWLHAPSLK 89
: | | | | |
DB 551 RPSAHS---HAGSLR 562
RESULT 12
AAM93401
ID AAM93401 standard; Protein; 740 AA.
XX
AC AAM93401;
XX
DT 06-NOV-2001 (first entry)
XX
```

```
XX
DE Human polypeptide, SEQ ID NO: 3002.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR N-PSDB; AAK94322.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 3002; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 740 AA;
Query Match 14.1%; Score 87; DB 22; Length 740;
Best Local Similarity 31.7%; Pred. No. 2.8;
Matches 33; Conservative 10; Mismatches 45; Indels 16; Gaps 4;
QY 9 ALSRRARRLAYMRRALESSQRPRGSPRRSQNNIYSACP-RRARGADAAGTGEAPVPGP 67
|||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 223 AVSPEPRPAPAVSPGSKWPKPGPGSPRPWKSNPSASSGPKWPKAPKPSVS-----PGPW 276
QY 68 APLPPAPLQVSESPWLHAPSL-----KTSCEYVSLYHQPAAMME 106
|||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 277 KPFP----SVSPGPKWPKTPSVSSASWKS SVSPSSWKSPPASPE 316
RESULT 13
ABG12999
ID ABG12999 standard; Protein; 343 AA.
XX
AC ABG12999;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #12990.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
```

```

PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI WPI: 2001-639362/73.
XX N-PSDB; AAS71186.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID No 43358; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 343 AA;
Query Match 14.0%; Score 86; DB 22; Length 343;
Best Local Similarity 40.0%; Pred. No. 1.5;
Matches 30; Conservative 6; Mismatches 31; Indels 8; Gaps 3;

Qy 13 RARRLAVRMRALESSQPRGSPRRSONNIYSACPRRAGADAAGTGEAPVPGGAPL-- 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 RGRCLSGRLRHQSSPQSGGSGMAGCRSR---AMPRR--AAKAGKSSPVPVGGTALST 56
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 71 -PPAPLVQVSESPWLH 84
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 57 GPPQPLAIVLSPSLH 71

RESULT 14
ID AAR79048
XX AAR79048 standard; Protein; 2205 AA.
AC
XX
XX AAR79048;
XX
XX 27-FEB-1996 (first entry)
DT
XX
XX Infectious rubella virus N-terminal transcript.
DE
XX
XX Rubella; vaccine; mutant; epitope; virus; autoimmune disease;
KW pregnancy; foetal infection; vector; plasmid.
XX
XX

```

```

OS Rubella virus.
XX
XX US5439814-A.
XX
XX 08-AUG-1995.
XX
XX 28-JUN-1991; 91US-0722334.
PF
XX
XX 28-JUN-1991; 91US-0722334.
PR 19-JUL-1993; 93US-0093453.
PR
XX
XX (GEOR-) GEORGIA STATE RES FOUND INC.
PA
XX
XX Dominguez G, Frey TK, Wang C;
PI
XX
XX WPI; 1995-283097/37.
DR N-PSDB; AAG97686.
DR
XX
XX New DNA encoding infectious rubella virus - esp. non-pathogenic
PT mutant virus for use in vaccines having reduced side effects
PT
XX
XX Disclosure; Columns 17-28; 27pp; English.
PS
XX
XX Non-pathogenic mutants of the DNA corresponding to this sequence are
CC useful in vaccines (which may include epitopes from other viruses).
CC The mutant vaccines are less likely to cause foetal infections,
CC autoimmune disease or neurological symptoms, so can be administered
CC safely to older or pregnant women. The mutant sequence is
CC preferably present in a vector, especially a bacterial plasmid that
CC allows replication of the sequence.
XX
XX SQ Sequence 2205 AA;
Query Match 13.9%; Score 85.5; DB 16; Length 2205;
Best Local Similarity 31.4%; Pred. No. 13;
Matches 38; Conservative 9; Mismatches 43; Indels 31; Gaps 7;

Qy 9 ALSRRRLAVRM--RALESSQPRGSPRRSONNIYSACPRRA---RGADAAGT---G 59
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 699 AFDANAVTAAVRAGPQSAASPPGDPDP-----PRRARSQRHSOARGTTPPPA 748
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 60 EAPVPGGAPLPPAPLVQVSE--SPWLHPS-----LKTCEYVSLYHQPAAAMMEDSDS 110
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 749 PARDPDPVPPAPPPAPPPAGDPVPPIPAGPADRADARDAELEVACEPSG---PPTSTRADPDS 805
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 111 D 111
   |
Db 806 D 806

RESULT 15
ID AAG90922
XX AAG90922 standard; Protein; 213 AA.
XX
XX AAG90922;
AC
XX
XX 26-SEP-2001 (first entry)
DT
XX
XX C glutamicum protein fragment SEQ ID NO: 4676.
DE
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
XX Corynebacterium glutamicum.
XX
XX EPI108790-A2.
PN
XX
XX 20-JUN-2001.
PD
XX
XX 18-DEC-2000; 2000EP-0127688.
PF
XX
XX 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR

```

```

PR 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
DR N-PSDB; AAH66141.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
XX Claim 17; SEQ ID NO: 4676; 246pp + Sequence Listing; English.
PS
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
XX Sequence 213 AA;
SQ
Query Match 13.88; Score 85; DB 22; Length 213;
Best Local Similarity 28.6%; Pred. No. 1.1;
Matches 28; Conservative 15; Mismatches 45; Indels 10; Gaps 3;
QY 4 VERKALSRARRLAVRMRALESORPRGSPRPRSONNIYSACPRRARGADAAGTGEAPV 63
: ||: ||| || : : : : : : : : : : : : : : : : : : : : : :
Db 13 LRRRLSLRSARSLAKVSAETRAASSPSWRRIASMSRSVWELSCSARLSSARSAMAPV 72
| | : | : | | | | | | | | | | | | | | | | | | | | | |
QY 64 PPGAPLPPAPLQVSE-----SPW-LHAPSLKTSC 92
| | : | : | | | | | | | | | | | | | | | | | | | | |
Db 73 PSP-SPMPSSPENCSEPSQLRSGRRPWRFPARSPNSC 109
| | : | : | | | | | | | | | | | | | | | | | | | | |
Search completed: October 28, 2002, 17:30:57
Job time : 15.371 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:28:11 : Search time 9.83333 Seconds
(without alignments)
586.214 Million cell updates/sec

Title: US-09-135-238B-2_COPY_18_253

Perfect score: 1254

Sequence: 1 RILPEVKVEGELGSGVTIKC.....HRQRALDYGSQSGREGQGF 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	390	3	US-08-961-564A-2
2	459	36.6	107	3	US-08-961-564A-4
3	197	15.7	771	3	US-08-434-000A-8
4	197	15.7	771	4	US-09-312-157-8
5	187	14.9	109	3	US-08-961-564A-9
6	185	14.8	769	3	US-08-434-000A-10
7	185	14.8	769	4	US-09-312-157-10
8	180.5	14.4	608	4	US-09-095-385-4
9	180.5	14.4	746	3	US-08-434-000A-4
10	180.5	14.4	746	4	US-09-312-157-4
11	179.5	14.3	757	3	US-08-434-000A-6
12	179.5	14.3	757	4	US-09-312-157-6
13	166.5	13.3	624	2	US-08-642-406A-22
14	166.5	13.3	624	4	US-09-199-534-22
15	166.5	13.3	773	3	US-08-434-000A-2
16	166.5	13.3	773	4	US-09-312-157-2
17	96.5	7.7	476	3	US-08-487-550-4
18	90	7.2	453	3	US-08-466-151-8
19	90	7.2	453	4	US-08-466-163B-8
20	89.5	7.1	451	2	US-08-887-352B-14
21	89.5	7.1	451	2	US-08-887-352B-16
22	89.5	7.1	451	3	US-08-466-151-65
23	89.5	7.1	451	4	US-09-109-207C-14
24	89.5	7.1	451	4	US-09-109-207C-16
25	89.5	7.1	451	4	US-09-296-005-14
26	89.5	7.1	451	4	US-09-296-005-16
-	86.5	6.9	46	3	US-08-955-937A-10

Sequence 10, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 18, Appl
Patent No. 520236
Sequence 4, Appl
Sequence 4, Appl
Sequence 12, Appl
Sequence 21, Appl
Sequence 13, Appl
Sequence 26, Appl
Sequence 4, Appl
Sequence 26, Appl
Sequence 27, Appl

28 86.5 6.9 46 4 US-09-300-985-10
29 85.5 6.8 229 2 US-08-887-352B-20
30 85.5 6.8 229 4 US-09-109-207C-20
31 85.5 6.8 229 4 US-09-296-005-20
32 85.5 6.8 233 2 US-08-887-352B-25
33 85.5 6.8 233 4 US-09-109-207C-25
34 85.5 6.8 233 4 US-09-296-005-25
35 84.5 6.7 87 3 US-08-554-840-18
36 84.5 6.7 744 6 520236-25
37 84 6.7 195 3 US-08-955-937A-4
38 84 6.7 195 4 US-09-300-985-4
39 84 6.7 275 2 US-08-511-485-12
40 84 6.7 275 3 US-08-836-134-21
41 84 6.7 449 1 US-08-458-516-13
42 83.5 6.7 461 2 US-08-463-587A-26
43 83.5 6.7 461 2 US-08-463-667A-4
44 83.5 6.7 461 3 US-08-923-854-26
45 83.5 6.7 461 5 PCT-US91-09133-27

ALIGNMENTS

RESULT 1

US-08-961-564A-2
; Sequence 2, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-2

Query Match 100.0%; Score 1254; DB 3; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.le-119;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 RILPEVKVEGELGSSVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTNFIKAEYGRVT 60
Db 18 RILPEVKVEGELGSSVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTNFIKAEYGRVT 77
QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSWEEOPMPE 120
Db 78 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSWEEOPMPE 137
QY 121 TPKWFHLPYLPOMPAYASSSKFVTRVTPAQRGKVPVHHSSPTTQIHRPRVSRASSVA 180
Db 138 TPKWFHLPYLPOMPAYASSSKFVTRVTPAQRGKVPVHHSSPTTQIHRPRVSRASSVA 197
QY 181 GDKPRTLPSTTASKISALEGLKLPQTPSYNHHHTLRLHQALDYGSGSGREGOGFH 236
Db 198 GDKPRTLPSTTASKISALEGLKLPQTPSYNHHHTLRLHQALDYGSGSGREGOGFH 253

RESULT 2

US-08-961-564A-4
; Sequence 4, Application US/08961564A

; Patent No. 6114515

; GENERAL INFORMATION:

; APPLICANT: WU, SHUJIAN

; APPLICANT: SWEET, RAYMOND

; APPLICANT: TRUNEH, ALEMSEGED

; TITLE OF INVENTION: FIGURE-1, A MEMBER OF IMMUNOGLOBULIN

; TITLE OF INVENTION: GENE SUPERFAMILY

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,564A

; FILING DATE: 30-OCT-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/056,935

; FILING DATE: 25-AUG-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-70236

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 107 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-961-564A-4

Query Match 36.6%; Score 459; DB 3; Length 107;

Best Local Similarity 100.0%; Pred. No. 1.7e-39;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILPEVKVEGELGSSVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTNFIKAEYGRVT 60
Db 18 RILPEVKVEGELGSSVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTNFIKAEYGRVT 77

QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACG 88
Db 78 LKQYPRKNLFLVEVTQLTESDSGVYACG 105

RESULT 3

US-08-434-000A-8

; Sequence 8, Application US/08434000A

; Patent No. 6046037

; GENERAL INFORMATION:

; APPLICANT: ANDREW C. HIATT, JULIAN

; APPLICANT: K.-C. MA, THOMAS LEHNER

; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION

; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; CITY: Suite 4700

; STATE: Los Angeles

; COUNTRY: California

; ZIP: U.S.A.

; ZIP: 90071

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: Word Perfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/434,000A

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; PRIOR APPLICATION DATA: including application

; PRIOR APPLICATION DATA: described below:

; APPLICATION NUMBER: 08/367,395

; FILING DATE: 12/30/94

; ATTORNEY/AGENT INFORMATION:

; NAME: Guise, Jeffrey W.

; REGISTRATION NUMBER: 34,613

; REFERENCE/DOCKET NUMBER: 212/127

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 552-8400

; TELEFAX: (619) 552-0159

; TELEX: 67-3510

; SEQUENCE LISTING

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 771 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; DESCRIPTION:

; US-08-434-000A-8

Query Match

Best Local Similarity 15.7%; Score 197; DB 3; Length 771;

Matches 48; Conservative 22; Mismatches 54; Indels 24; Gaps 5;

QY 13 GGSVTIKCPLPE----MHVRIYLCREMAGSGTCGTVSTNFIKAEYGRVTLKQYPRKN 68
Db 33 GDSVITCYYPDTSVNRHTRKYWCROGA-SGMCTTLISSNGYLSKEYSGRANLINFPENN 91
QY 69 LFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSWEEOPMETPKWFHLP 128
Db 92 TFVINIEQTQDTSYKCGCLG-TSNRGLSFDVSLEV-----SQVPELPSDTHV- 139

QY 129 YLFQMPAYASSSKFVTRVTPAQRGKVP 156
Db 140 -----YTKDIGRNVITBCPFKRENVP 160

1

2

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:28:11 ; Search time 16.25 Seconds
(without alignments)
586.214 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLWPLFLPVSGALRIL.....HQPAAMWEDSDDDYINVPA 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	99.6	390	3	US-08-961-564A-2
2	550	26.8	107	3	US-08-961-564A-4
3	198	9.6	771	3	US-08-434-000A-8
4	198	9.6	771	4	US-09-312-157-8
5	187	9.1	109	3	US-08-961-564A-9
6	185	9.0	769	3	US-08-434-000A-10
7	185	9.0	769	4	US-09-312-157-10
8	182	8.9	608	4	US-09-095-385-4
9	182	8.9	746	3	US-08-434-000A-4
10	182	8.9	746	4	US-09-312-157-4
11	182	8.9	757	3	US-08-434-000A-6
12	182	8.9	757	4	US-09-312-157-6
13	166.5	8.1	624	2	US-08-642-406A-22
14	166.5	8.1	624	4	US-09-199-534-22
15	166.5	8.1	773	3	US-08-434-000A-2
16	166.5	8.1	773	4	US-09-312-157-2
17	117.5	5.7	476	3	US-08-487-550-4
18	117	5.7	453	3	US-08-466-151-8
19	117	5.7	453	4	US-08-466-163B-8
20	113.5	5.5	2337	3	US-08-713-118-2
21	113.5	5.5	2337	4	US-09-452-007-2
22	112	5.5	451	2	US-08-887-352B-14
23	112	5.5	451	2	US-08-887-352B-16
24	112	5.5	451	3	US-08-466-151-65
25	112	5.5	451	4	US-09-109-207C-14
26	112	5.5	451	4	US-09-109-207C-16
27	112	5.5	451	4	US-09-296-005-14

28 112 5.5 451 4 US-09-296-005-16 Sequence 16, Appl
29 111 5.4 476 2 US-08-378-939-10 Sequence 10, Appl
30 109.5 5.3 2339 1 US-08-455-543A-47 Sequence 47, Appl
31 109.5 5.3 2339 2 US-08-223-305C-47 Sequence 47, Appl
32 109.5 5.3 2339 4 US-09-268-163-6 Sequence 6, Appl
33 109.5 5.3 2343 4 US-09-268-163-4 Sequence 4, Appl
34 107 5.2 292 4 US-09-345-468-18 Sequence 18, Appl
35 107 5.2 313 4 US-09-345-468-16 Sequence 16, Appl
36 107 5.2 449 1 US-08-458-516-13 Sequence 13, Appl
37 106 5.2 390 2 US-08-979-424-1 Sequence 1, Appl
38 105 5.1 451 2 US-08-887-352B-18 Sequence 18, Appl
39 105 5.1 451 4 US-09-109-207C-18 Sequence 18, Appl
40 105 5.1 451 4 US-09-282-505-2 Sequence 2, Appl
41 105 5.1 451 4 US-09-054-255-2 Sequence 2, Appl
42 105 5.1 451 4 US-09-296-005-18 Sequence 18, Appl
43 104 5.1 554 1 US-08-347-254-1 Sequence 1, Appl
44 104 5.1 554 2 US-08-464-463-1 Sequence 1, Appl
45 103.5 5.0 459 1 US-08-157-101A-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-961-564A-2
; Sequence 2, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-2

Query Match 99.6%; Score 2047; DB 3; Length 390;
Best Local Similarity 99.7%; Pred. No. 2e-178;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



QY 1 MDRWLWPLFLPVSGALRILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGT 60
DB 1 MDRWLWPLFLPVSGALRILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGT 60
QY 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQJTESDGYVACGAGMNTDRGKTQVTL 120
DB 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQJTESDGYVACGAGMNTDRGKTQVTL 120
QY 121 NVHSEYEFSEWEOQMPETPKFHLPLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSP 180
DB 121 NVHSEYEFSEWEOQMPETPKFHLPLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSP 180
QY 181 TQIITHRPRVSRASSVAGDKPRTLPSTASKISALEGLLKPQTPSYNHHTRLHRQALD 240
DB 181 TQIITHRPRVSRASSVAGDKPRTLPSTASKISALEGLLKPQTPSYNHHTRLHRQALD 240
QY 241 YGSQSGRGQGFHILPILGLFLALLGLVVKRAVERRKALSRARRLAVRMALESSQ 300
DB 241 YGSQSGRGQGFHILPILGLFLALLGLVVKRAVERRKALSRARRLAVRMALESSQ 300
QY 301 RPRGSPRPRSONNIYSACPRRARGAAGTGEAPVPGGAPLPAPLOVSESPWLHAPSL 360
DB 301 RPRGSPRPRSONNIYSACPRRARGAAGTGEAPVPGGAPLPAPLOVSESPWLHAPSL 360
QY 361 KTSCEYVSLYHOPAAAMMEDSDDDYINVPA 390
DB 361 KTSCEYVSLYHOPAAAMMEDSDDDYINVPA 390

RESULT 2

US-08-961-564A-4
; Sequence 4, Application US/08961564A
; Patent No. 6114515

GENERAL INFORMATION:

APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND
APPLICANT: TRUNEH, ALEMSEGED
TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
TITLE OF INVENTION: GENE SUPERFAMILY
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/08/961,564A
FILING DATE: 30-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/056,935
FILING DATE: 25-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70236
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-564A-4

Query Match 26.8%; Score 550; DB 3; Length 107;
Best Local Similarity 99.0%; Pred. No. 6.6e-43;
Matches 104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLWPLFLPVSGALRILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGT 60
DB 1 MDRWLWPLFLPVSGALRILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGT 60
QY 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQJTESDGYVACG 105
DB 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQJTESDGYVACG 105

RESULT 3

US-08-434-000A-8
; Sequence 8, Application US/08434000A
; Patent No. 6046037

GENERAL INFORMATION:

APPLICANT: ANDREW C. HIATT, JULIAN
APPLICANT: K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,000A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 1
APPLICATION NUMBER: 08/367,395
FILING DATE: 12/30/94
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-3510
FILING DATE: 25-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70236
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
TOPOLOGY: DESCRIPTION: Mouse Polyimmunoglobulin Receptor
US-08-434-000A-8

Query Match 9.6%; Score 198; DB 3; Length 771;
Best Local Similarity 31.5%; Pred. No. 1.2e-09;
Matches 53; Conservative 23; Mismatches 66; Indels 24; Gaps 5;

QY 10 FLPVSGALRILPEVKVEGELGSGVTIKCPLPE----MHVRIYLCREMAGSGTCGTVSTT 65



GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 21.2527 Seconds
(without alignments)
1921.018 Million cell updates/sec

Title: US-09-135-238B-2_COPY_18_253
Perfect score: 1254
Sequence: 1 RILPEVKVGEIGSGVTIKC.....HRQRALDYGSQSGREGQGFH 236

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTRMBL_19.*
- 1: sp-archaea.*
 - 2: sp-bacteria.*
 - 3: sp-fungi.*
 - 4: sp-human.*
 - 5: sp-invertebrate.*
 - 6: sp-mammal.*
 - 7: sp-mhc.*
 - 8: sp-organelle.*
 - 9: sp-phage.*
 - 10: sp-plant.*
 - 11: sp-rodent.*
 - 12: sp-virus.*
 - 13: sp-vertebrate.*
 - 14: sp-unclassified.*
 - 15: sp-rvirus.*
 - 16: sp-bacteriap.*
 - 17: sp-archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	390	4	O60667
2	654	52.2	422	11	Q9D8T1
3	181	14.4	758	6	Q9N2H7
4	175.5	14.0	534	4	Q96SA2
5	166	13.2	455	11	Q920L8
6	166	13.2	535	11	Q9EQT7
7	121	9.6	299	4	Q9UBK4
8	121	9.6	299	4	Q9UGN4
9	117.5	9.4	359	4	O43656
10	111.5	8.9	307	11	O34947
11	108	8.6	364	4	Q96D42
12	106	8.5	298	4	Q9HD97
13	106	8.5	301	4	O95100
14	105	8.4	270	4	Q9DWT1
15	105	8.4	270	4	Q9H564
16	100.5	8.0	503	4	Q9H8V0

17	100.5	8.0	504	4	Q9UD78	Q9ud78 homo sapien
18	100	8.0	227	11	Q9JL34	Q9jl34 mus musculu
19	98.5	7.9	230	4	Q9NZC2	Q9nzc2 homo sapien
20	98	7.8	886	13	P70049	P70049 xenopus lae
21	97.5	7.8	662	4	O60926	O60926 homo sapien
22	96	7.7	230	11	Q9JKE2	Q9jke2 mus musculu
23	96	7.7	589	4	Q9H802	Q9h802 homo sapien
24	95	7.6	258	4	Q9DWT2	Q9dwt2 homo sapien
25	95	7.6	258	4	Q9H563	Q9h563 mus musculu
26	94	7.5	227	11	Q9NNH8	Q9nnh8 mus musculu
27	93.5	7.5	1847	5	Q9NKN5	Q9nkn5 leishmania
28	92.5	7.4	336	13	Q90Z89	Q90z89 brachydanio
29	92.5	7.4	504	11	O60786	O60786 mus musculu
30	92	7.3	287	11	Q9D7B8	Q9d7b8 mus musculu
31	91.5	7.3	502	4	Q9UD77	Q9ud77 homo sapien
32	91.5	7.3	502	4	Q12800	Q12800 homo sapien
33	91.5	7.3	4072	5	Q9W4Y4	Q9w4y4 drosophila
34	91	7.3	227	11	Q9NNH9	Q9nnh9 mus musculu
35	91	7.3	757	5	Q9V6T1	Q9v6t1 drosophila
36	90.5	7.2	315	4	O96K52	O96k52 homo sapien
37	90.5	7.2	549	5	Q9U3F1	Q9u3f1 caenorhabdi
38	90.5	7.2	553	5	Q9GTC3	Q9gtc3 caenorhabdi
39	90.5	7.2	556	5	Q9XXU7	Q9xxu7 caenorhabdi
40	90.5	7.2	614	4	O96FE5	O96fe5 homo sapien
41	90.5	7.2	614	11	Q9D1T0	Q9d1t0 mus musculu
42	90.5	7.2	754	4	O96KV7	O96kv7 homo sapien
43	90.5	7.2	1209	11	Q924W6	Q924w6 mus musculu
44	90	7.2	822	5	O61674	O61674 drosophila
45	89.5	7.1	1417	12	O67631	O67631 gallid herp

ALIGNMENTS

RESULT 1

O60667 ID O60667 PRELIMINARY; PRT; 390 AA.
AC O60667;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ANTI-FAS-INDUCED APOPTOSIS (REGULATOR OF FAS-INDUCED APOPTOSIS).
GN TOSO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98245048; PubMed=9586636;
RA Hitoshi Y., Lorens J., Kitada S.I., Fisher J., Labarge M., Ring H.Z.,
RA Francke U., Reed J.C., Kinoshita S., Nolan G.P.;
RT "Toso, a cell surface, specific regulator of Fas-induced apoptosis in
T cells";
RL Immunity 8:461-471(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF057557; AAC18830.1; -;
DR EMBL; BC006401; AAH06401.1; -;
DR InterPro; IPR003599; Ig_
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
SQ SEQUENCE 390 AA; 43146 MW; FE91D217EECA99C6 CRC64;

Query Match 100.0%; Score 1254; DB 4; Length 390;

Best Local Similarity 100.0%; Pred. No. 1.7e-115;

Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RILPEVKVGEIGSGVTIKCPLPEMHVRIYLCREMAGSGTCGVTVVTFNFIKAEYKGRVT 60

```

Db 18 RILPEVKVEELGGVTKCPLEPMHVRVILCREMAGSGTCGTVTTFIKAEYKGRVT 77
QY 61 LKQYPRKNLFLVEVQTLESQGYACGAGMNTDRGKTQKVTNLNVHSEYEPWESQPMPE 120
Db 78 LKQYPRKNLFLVEVQTLESQGYACGAGMNTDRGKTQKVTNLNVHSEYEPWESQPMPE 137
QY 121 TPKEPHLPYLFOFQPAVYASSKFVTRVTPAQRGKVPVPHHSSPTQITHRPRVRSASSVA 180
Db 138 TPKEPHLPYLFOFQPAVYASSKFVTRVTPAQRGKVPVPHHSSPTQITHRPRVRSASSVA 197
QY 181 GDKPRTFLPTTASIKALEGLLKQPSPSYNHHTRLHQRALDYGSGSQSGREGQGFH 236
Db 198 GDKPRTFLPTTASIKALEGLLKQPSPSYNHHTRLHQRALDYGSGSQSGREGQGFH 253

RESULT 2
Q9DBT1 PRELIMINARY; PRT; 422 AA.
AC Q9DBT1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1810037B05RIK PROTEIN.
GN 1810037B05RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
DR EMBL: AK007714; BAB25207.1; -.
DR MGI: 1916419; 1810037B05RIK.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: Pf00047; Ig; 1.
DR SMART: SM00410; Ig_like; 1.
SQ SEQUENCE 422 AA; 47532 MW; 2597083A50AD8B6E CRC64;

```

Query Match 52.2%; Score 654; DB 11; Length 422;
 Best Local Similarity 53.3%; Pred. No. 3.2e-56;
 Matches 129; Conservative 34; Mismatches 71; Indels 8; Gaps 4;

```

QY 1 RILPEVKVEELGGVTKCPLEPMHVRVILCREMAGSGTCGTVTTFIKAEYKGRVT 60
Db 18 RVLPEVLNVEGSGIIIECPQLQLHVRMYLCRQMAKPGICSTVVSNT-FVKKEYERRVT 76
QY 61 LKQYPRKNLFLVEVQTLESQGYACGAGMNTDRGKTQKVTNLNVHSEY-EPWESQPMPE 119
Db 77 LTPCLDKLFLVEMTQLENDGDIYACGVGMKTKGKTQKITLNVHNEYPEPFWDEWTS 136

```

```

QY 120 ETPKWFHLPYLFQMP-----AYASSKFVTRVTPAQRGKVPVPHHSSPTQITHRPRVS 174
Db 137 ERPNLHRFLQHPWLGSEHPSGSGVIKAVTTPPKTEAPPVHPQSSITSVTQHPRV 196
QY 175 RASVAGDKPRTFLPTTASIKALEGLLKQPSPSYNHHTRLHQRALDYGSGSQSGREGQ 234
Db 197 RAFSVSATKSPALLPATTTASKTSTQQA-IRPLEASYSHTLRLHQRHHGPHYGREDRG 255
QY 235 FH 236
Db 256 LH 257

RESULT 3
Q9N2H7 PRELIMINARY; PRT; 758 AA.
AC Q9N2H7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE POLY-IG RECEPTOR PRECURSOR.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN 1
SEQUENCE FROM N.A.
RA Sone T., Kumura H.;
RT "Porcine mammary gland cDNA clone, similar to poly-Ig receptor.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032195; BAA84283.2; -.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: Pf00047; Ig; 5.
DR SMART: SM00409; IG; 3.
DR SMART: SM00410; Ig_like; 2.
KW Signal; Receptor.
FT SIGNAL 1 18
FT SIGNAL 18
SQ SEQUENCE 758 AA; 83154 MW; D5BEBIA8B082D247 CRC64;

Query Match 14.4%; Score 181; DB 6; Length 758;
Best Local Similarity 33.9%; Pred. No. 3.4e-09;
Matches 40; Conservative 22; Mismatches 40; Indels 16; Gaps 4;

QY 2 ILPEVKVEGL-----GGSVTIKPLP-----EMHVRVILCREMAGSGTCGTVYST 47
Db 12 IFPVVSMKSPIFGPQDVSSVEGSSVRCYYPATSVNHRHRSKYWCR-IGAKGRCTLISS 70
QY 48 TNFIKAEYKGRVTLKQYPRKNLFLVEVQTLESQGYACGAGMNTDRGKTQKVTNLN 105
Db 71 EGYISKDYKGRANLTFNPENGTFFVMDIGHLTRGDSGLYKGLGISS-RGLSPFDSLEV 127

RESULT 4
Q96SA2 PRELIMINARY; PRT; 534 AA.
AC Q96SA2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FKSG87 PROTEIN.
GN FKSG87.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
SEQUENCE FROM N.A.
RA Wang Y.-G., Gong L.;
RT "Molecular cloning and characterization of FKSG87, a novel gene located on human chromosome 1.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

```

DR ENBL: AF354295; AAK39522.1; -.
SQ SEQUENCE 534 AA; 56748 MW; 6EF8050E412AF91C CRC64;

Query Match 14.0%; Score 175.5; DB 4; Length 534;
Best Local Similarity 21.9%; Pred. No. 7.5e-09;
Matches 67; Conservative 28; Mismatches 82; Indels 129; Gaps 10;

QY 8 VEGLGGSVTIKCPLP-----EMHVRIYLCRMAGSGTCGTVVSTTNFIKAEYKGRVTLKQ 63
Db 90 VSGEPGAVTIQCHYAPSSVNRHQRKYWCRLGPPRWICQTIVSTNOYTHRVDRVALTD 149
QY 64 YPRKNLFLEVTLQTESDGVYACAGMNTDR----- 95
Db 150 FPORGUFVRLSQLSPDDIGCYLGIGSENNMLFLSNLITISAGPASTLPTATPAAGELT 209
QY 96 -----GKTT-----KVTLVNHSEYEP----- 111
Db 210 MRSYGTASPVANRWTPGTTQTLGCGTAMDVTASTPGTSKTTASAEGRRTPGATRPAAQCT 269
QY 112 -SWE-----OPMETPKWHLPLYLFOMPAYASSKSFVTRVTPAQRGKVPVPHHSPT 164
Db 270 GSAEWSGVKAPAPIESP-----PSKRSMSNTTEGVWEG-----TR 306
QY 165 TQTHRPVRSR-----ASSVAGDKPR-----TELPSTTASKISALEGL 202
Db 307 SVYTNARASKDRREMTTKADRPREDIEGVRIALDAAKKVLGTIGPPALVSETLAWE-I 365
QY 203 LKPQTP 208
Db 366 LQQATP 371

RESULT 5
Q920L8 PRELIMINARY; PRT; 455 AA.
AC Q920L8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FCA/M RECEPTOR (FRAGMENT).
GN FCMR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu Y., Honda S., Yotsumoto K., Tahara-Hanaoka S., Eyre H.J.,
RA Sutherland G.D., Endo Y., Shibuya K., Koyama A., Nakauchi H.,
RA Shibuya A.;
RT "Fca/m receptor is a single gene-family member closely related to
RT polymeric immunoglobulin receptor on chromosome 1.";
RL Immunogenetics 0:0-0(2001).
DR ENBL: AB071978; BAB1750.1; -.
KW Receptor.
FT NON_TER
SQ SEQUENCE 455 AA; 48810 MW; 628913C334AC365 CRC64;

Query Match 13.2%; Score 166; DB 11; Length 455;
Best Local Similarity 28.0%; Pred. No. 5.3e-08;
Matches 65; Conservative 21; Mismatches 94; Indels 52; Gaps 8;

QY 8 VEGLGGSVTIKCPLP-----EMHVRIYLCRMAGSGTCGTVVSTTNFIKAEYKGRVTLKQ 63
Db 10 VTGNTGAVTIICHYAPSSVNRHQRKYWCRLGSLPWICTHTVSTNQTHDPYGRALTD 69
QY 64 YPRKNLFLEVTLQTESDGVYACAGMNTDRGKTQKVTLVNHSEYEPS---WEEQPMPE 120
Db 70 VPOSGLFVVRLRLSLGDLGYRCIG---DRNDMLFESVNLTVSAGPSTVTAAPASS 126
QY 121 TPKNFHLPLYLFOMPAYASSKSFVTRVTPAQRGKVPVPHHSPTQTQTHRPVRSASSVA 180
Db 127 EP-----TTASPGAASS-----AGNG-----WTSGVTOILEG-----S 154


```
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00410; Ig_like; 1.
SQ SEQUENCE 298 AA; 33230 MW; DB58C9104A33B4C3 CRC64;

Query Match      8.5%; Score 106; DB 4; Length 301;
Best Local Similarity 23.3%; Pred. No. 0.026;
Matches 47; Conservative 26; Mismatches 79; Indels 50; Gaps 10;

QY 14 GSVTIKCPLEMH--VRIYLCREMGSGTCGVVSTNFIKAEYKGRVTLKQVPRKNLFL 71
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 29 GSLSVQCPYQEHRTLNKYWCRP-PQIFLCKIVETKG-SAGKRNGRVSIKDSFANLSFT 86
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 72 VEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSEWEEQPMPTPKW---FHL 128
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 87 VTLENLTEDAGTWCGV-----DTP-WLRDFHDP 115
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 129 YL-FQMPAYASSKFEVTRVTPAQRKV-----PPVHSSPTTQI-----THRPVRV 179
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 116 VVEVEVSVPASTSMTPASTIAKTSTITTAFFPV---SSTLFAVGATHSASIQETEE 172
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 180 AGDKPRTPLPSTTASKISALEG 201
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 173 VVNSQLPLLLSLLALLLLLVG 194
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 13
ID O95100 PRELIMINARY; PRT; 301 AA.
AC O95100;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CMRF-35-H9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Green B.J., Clark G.J., Hart D.N.J.;
RT "The CMRF-35 monoclonal antibody recognised a second leucocyte
RL membrane molecule with a domain similar to the poly Ig receptor.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF020314; AAD01646.1; -.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00410; Ig_like; 1.
SQ SEQUENCE 301 AA; 33658 MW; 632059938BA48CF2 CRC64;
```

```
Query Match      8.5%; Score 106; DB 4; Length 301;
Best Local Similarity 23.3%; Pred. No. 0.026;
Matches 47; Conservative 26; Mismatches 79; Indels 50; Gaps 10;

QY 14 GSVTIKCPLEMH--VRIYLCREMGSGTCGVVSTNFIKAEYKGRVTLKQVPRKNLFL 71
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 29 GSLSVQCPYQEHRTLNKYWCRP-PQIFLCKIVETKG-SAGKRNGRVSIKDSFANLSFT 86
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 72 VEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSEWEEQPMPTPKW---FHL 128
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 87 VTLENLTEDAGTWCGV-----DTP-WLRDFHDP 115
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 129 YL-FQMPAYASSKFEVTRVTPAQRKV-----PPVHSSPTTQI-----THRPVRV 179
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 116 VVEVEVSVPASTSMTPASTIAKTSTITTAFFPV---SSTLFAVGATHSASIQETEE 172
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 180 AGDKPRTPLPSTTASKISALEG 201
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 173 VVNSQLPLLLSLLALLLLLVG 194
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```
RESULT 14
Q9UMT1 PRELIMINARY; PRT; 270 AA.
ID Q9UMT1;
AC Q9UMT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NKP44RG2.
DE NKP44RG2.
GN NKP44RG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOID;
RA Cantoni C., Biassoni R.;
RT "NKP44 related genes.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ010100; CAB52290.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.
SQ SEQUENCE 270 AA; 29678 MW; 327AD57A5634AE46 CRC64;

Query Match      8.4%; Score 105; DB 4; Length 270;
Best Local Similarity 21.8%; Pred. No. 0.028;
Matches 58; Conservative 31; Mismatches 97; Indels 80; Gaps 11;

QY 13 GGSVTIKCPPLP---EMHVRILCREMGSGTCGVVSTNFIKAEYKGRVTLKQVPRKNL 69
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 33 GQTLFVRCQPTPTGSLYEKKGWCKE-ASALVCLRLVTSSKPTWANTSRITWDDEDA 91
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 70 FLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSEWEEQPMPTPKWPHLP 129
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 92 FVTWTLDRSDSGHYWCRIYRPSDMSVSRF-----Y 126
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 130 LFQMPAYAS-----SSKFVTRVTPAQR-----KVP-PVHSSP-TTQI 167
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 127 LVVSPASASTQTPTWTRPDLVSSQTQSCVPTAGARQAPESPSTIPVSPSPPLPVPL 186
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 168 THRPRVS--RASSVAGDKPRTFLP---STTAKISALEGL----- 203
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 187 PSRPQNSTLRPGPAA---PIALVPVFCGLLVAKSLVLSALLVWVYLRNRHMQHGRSL 243
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 204 --KPQTPSYNHHTLHRQALDYGSO 227
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 244 PAQPRQAHRHPLSHRAPGGTYGK 269
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 15
Q9H564 PRELIMINARY; PRT; 270 AA.
ID Q9H564;
AC Q9H564;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DJ149M18.1.2 (NATURAL KILLER CELL P44-RELATED GENE 2 (NKP44RG2)).
GN DJ149M18.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136967; CAC09452.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_like.
```


GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:28:11 ; Search time 9.83333 Seconds
(without alignments)
586.214 Million cell updates/sec

Title: US-09-135-238B-2_COPY_18_253
Perfect score: 1254
Sequence: 1 RILPEVKVEGLGGVTKIC.....HRQRALDYGSOSGRGQGFH 236

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.5/pdata/1/iaa/5A_COMB.pep:*
2: /cgn2.5/pdata/1/iaa/5B_COMB.pep:*
3: /cgn2.5/pdata/1/iaa/6A_COMB.pep:*
4: /cgn2.5/pdata/1/iaa/6B_COMB.pep:*
5: /cgn2.5/pdata/1/iaa/PCUS_COMB.pep:*
6: /cgn2.5/pdata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	390	3 US-08-961-564A-2	Sequence 2, Appli
2	459	36.6	107	3 US-08-961-564A-4	Sequence 4, Appli
3	197	15.7	771	3 US-08-434-000A-8	Sequence 8, Appli
4	197	15.7	771	3 US-09-312-157-8	Sequence 8, Appli
5	187	14.9	109	3 US-08-961-564A-9	Sequence 9, Appli
6	185	14.8	769	3 US-08-434-000A-10	Sequence 10, Appl
7	185	14.8	769	4 US-09-312-157-10	Sequence 10, Appl
8	180.5	14.4	608	4 US-09-095-385-4	Sequence 4, Appli
9	180.5	14.4	746	3 US-08-434-000A-4	Sequence 4, Appli
10	180.5	14.4	746	3 US-09-312-157-4	Sequence 4, Appli
11	179.5	14.3	757	3 US-08-434-000A-6	Sequence 6, Appli
12	179.5	14.3	757	4 US-09-312-157-6	Sequence 6, Appli
13	166.5	13.3	624	2 US-08-642-406A-22	Sequence 22, Appl
14	166.5	13.3	624	4 US-09-199-534-22	Sequence 22, Appl
15	166.5	13.3	773	3 US-08-434-000A-2	Sequence 2, Appli
16	166.5	13.3	773	4 US-09-312-157-2	Sequence 2, Appli
17	96.5	7.7	476	3 US-08-487-550-4	Sequence 4, Appli
18	90	7.2	453	3 US-08-466-151-8	Sequence 8, Appli
19	90	7.2	453	4 US-08-466-163B-8	Sequence 8, Appli
20	89.5	7.1	451	2 US-08-887-352B-14	Sequence 14, Appl
21	89.5	7.1	451	2 US-08-887-352B-16	Sequence 16, Appl
22	89.5	7.1	451	3 US-08-466-151-65	Sequence 65, Appl
23	89.5	7.1	451	4 US-09-109-207C-14	Sequence 14, Appl
24	89.5	7.1	451	4 US-09-109-207C-16	Sequence 16, Appl
25	89.5	7.1	451	4 US-09-296-005-14	Sequence 14, Appl
26	89.5	7.1	451	4 US-09-296-005-16	Sequence 16, Appl
27	86.5	6.9	46	3 US-08-955-937A-10	Sequence 10, Appl

28	86.5	6.9	46	4 US-09-300-985-10	Sequence 10, Appl
29	85.5	6.8	229	2 US-08-887-352B-20	Sequence 20, Appl
30	85.5	6.8	229	4 US-09-109-207C-20	Sequence 20, Appl
31	85.5	6.8	229	4 US-09-296-005-20	Sequence 20, Appl
32	85.5	6.8	233	2 US-08-887-352B-25	Sequence 25, Appl
33	85.5	6.8	233	4 US-09-109-207C-25	Sequence 25, Appl
34	85.5	6.8	233	4 US-09-296-005-25	Sequence 25, Appl
35	84.5	6.7	87	3 US-08-554-840-18	Sequence 18, Appl
36	84.5	6.7	744	6 5202236-25	Patent No. 5202236
37	84	6.7	195	3 US-08-955-937A-4	Sequence 4, Appli
38	84	6.7	195	4 US-09-300-985-4	Sequence 4, Appli
39	84	6.7	275	2 US-08-511-485-12	Sequence 12, Appl
40	84	6.7	275	3 US-08-836-134-21	Sequence 21, Appl
41	84	6.7	449	1 US-08-458-516-13	Sequence 13, Appl
42	83.5	6.7	461	2 US-08-463-587A-26	Sequence 26, Appl
43	83.5	6.7	461	2 US-08-463-667A-4	Sequence 4, Appli
44	83.5	6.7	461	3 US-08-923-854-26	Sequence 26, Appl
45	83.5	6.7	461	5 PCT-US91-09133-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-08-961-564A-2
; Sequence 2, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; TITLE OF INVENTION: TRUNEH, ALEMGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-2

Query Match 100.0%; Score 1254; DB 3; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.le-119;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 1
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 769 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Rat Polyimmunoglobulin Receptor
; US-08-434-000A-10

Query Match 14.8%; Score 185; DB 3; Length 769;
Best Local Similarity 34.5%; Pred. No. 2.2e-10;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 13 GGSVTIKPLPE----MHVRYILCREMAGSGTGTGVVSTTFNFIKAEYKGRVTLKQYPRKN 68
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 33 GNSVITCYPTDSVNRHTRKYWCROGA-NGYCATLISSNGVLSKEYSGRASLINFPENS 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 69 LFLVEVTLTSDSGVYACGAGMNTDRGKTQKVTNLNVHSEYSPWEEQPMPTPKWFHL 127
| : : | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 92 TFVINAHLTQEDTGSYKCGLG-TTNRGLFFDVSLEV-----SQVPEPPNDTHV 139
| : : | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 7
US-09-312-157-10
; Sequence 10, Application US/09312157
; Patent No. 6303341
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,157
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,000
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-351
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 769 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Rat Polyimmunoglobulin Receptor
; US-09-312-157-10

Query Match 14.8%; Score 185; DB 4; Length 769;
Best Local Similarity 34.5%; Pred. No. 2.2e-10;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 13 GGSVTIKPLPE----MHVRYILCREMAGSGTGTGVVSTTFNFIKAEYKGRVTLKQYPRKN 68
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 33 GNSVITCYPTDSVNRHTRKYWCROGA-NGYCATLISSNGVLSKEYSGRASLINFPENS 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 69 LFLVEVTLTSDSGVYACGAGMNTDRGKTQKVTNLNVHSEYSPWEEQPMPTPKWFHL 127
| : : | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 92 TFVINAHLTQEDTGSYKCGLG-TTNRGLFFDVSLEV-----SQVPEPPNDTHV 139
| : : | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 8
US-09-095-385-4
; Sequence 4, Application US/09095385
; Patent No. 6300104
; GENERAL INFORMATION:
; APPLICANT: Morrison, Sherie L.
; APPLICANT: Chintalacharuvu, Kote R.
; TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED
; BY SINGLE CELLS AND METHODS FOR MAKING AND USING
; TITLE OF INVENTION: SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 11150 Santa Monica Boulevard, Suite 400
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,385
; FILING DATE: 09-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,969
; FILING DATE: 19-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Canady, Karen S
; REGISTRATION NUMBER: 39,927
; REFERENCE/DOCKET NUMBER: 30435.45USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310 445-1140
; TELEFAX: 310 445-9031
; TELEX:
```

```

; INFORMATION FOR SEQ ID NO: 4:
;-----
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 608 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: unknown
;     TOPOLOGY: linear
;     MOLTYPE: protein
;     US-09-095-385-4

```

Query Match 14.4%; Score 180.5; DB 4; Length 608;
Best Local Similarity 38.7%; Pred. NO. 4.5e-10;
Matches 41; Conservative 19; Mismatches 39; Indels 7

QY	4	PEWKVEGELGSGVTTIKCLP----	EMHVRVYLCREMGAGSTCTGVVSTTNFNTKAEYKGRV	59
Db	25	PE-EVNSVEGNSVTCYPTSVNRHTRKYWCQARGG-CITLISSEGYSSKYAGRA	82	
QY	60	TLKQYPRKNLFVETQLTSDSGVYACGAGMNTDRGKTOKVTLNV	105	
Db	83	NLTNPENCTFVNIAOLSDDSGRYKCGLGINS-RCGLSPDVSLEV	127	

```

: RESULT 9
: US-08-434-000A-4
: Sequence 4, Application US/08434000A
: Patent No. 6046037
: GENERAL INFORMATION:
: APPLICANT: ANDREW C. HIATT, JULIAN
: APPLICANT: K. -C. MA, THOMAS LEHNER
: TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
: TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES

```

Query Match 14.4%; Score 180.5; DB 3; Length 746;
Best Local Similarity 38.7%; Pred. No. 6e-10;
Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;

59	Qy	4	PEWKVEGELGSGVTTIKCP	-----EMHVRIVYLREMA	SGCTCTVYSTNFTNFKAEYKGRV	59
60	Qy	5	PE-EVNSVEGNSVTCYPP	TSVNRHFRKWKCRQGRAGG	-CITLSSEGYVSSSYAGRA	64
61	Db	6	TLKQYPRKRLFLVEV	TLQTSDSGYACGAGMNTDRGKTQKTVLNV	105	
62	Qy	7	NLTNPENGTFFVNTAQL	SDQSRRYKCGLGINS	-RGLSFDVSLV	109
63	Db	8	TLKQYPRKRLFLVEV	TLQTSDSGYACGAGMNTDRGKTQKTVLNV	105	

```

: RESULT 10
: US-09-312-157-4
: Sequence 4, Application US/09312157
: Patent No. 6303341
: GENERAL INFORMATION:
: APPLICANT: ANDREW C. HIATT, JULIAN
: K.-C. MA, THOMAS LEHNER
: TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
: PROTEINS IN PLANTS AND THEIR USES
:

```

Query Match	14.4%;	Score 180.5;	DB 4;	Length 746;
Best Local Similarity	38.7%;	Pred. No. 6e-10;		
Matches 41: Conservative	19;	Mismatches 39;	Indels 7;	Gaps 4;

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 10.6263 seconds
(without alignments)
1921.018 Million cell updates/sec

Title: US-09-135-238B-2_COPY_273_390
Perfect score: 616
Sequence: 1 KRAVERKALSRARRLAVR.....HQPAAEMDSDDYINVPA 118

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_protist:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rviris:*

16: sp_bacteriaph:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	616	100.0	390	4	O60667
2	361.5	58.7	422	11	Q9D8T1
3	96	15.6	114	11	Q64561
4	96	15.6	115	11	Q64390
5	93	15.1	559	5	O46132
6	89.5	14.5	163	12	O9WSV8
7	89.5	14.5	748	10	Q9AH83
8	88.5	14.4	526	16	Q9RUB1
9	87.5	14.2	2936	5	Q9NKP7
10	87	14.1	821	4	Q96JM3
11	85.5	13.9	166	2	O30801
12	85	13.8	523	5	O77238
13	85	13.8	554	5	O9VTX7
14	84.5	13.7	152	12	Q9JH32
15	84.5	13.7	219	4	Q96PJ4
16	84.5	13.7	251	2	O9EY22

17	84.5	13.7	359	4	Q9BR57
18	84.5	13.7	1049	10	Q9ZNR9
19	84.5	13.7	2129	12	Q9J6K9
20	84	13.6	384	11	Q9CUF6
21	83	13.5	161	2	Q9F180
22	83	13.5	706	5	Q9VYR5
23	83	13.5	905	11	Q9WW47
24	83	13.5	1039	2	Q93L97
25	83	13.5	1260	5	Q9V468
26	83	13.5	1740	11	Q9JLU4
27	83	13.5	1806	11	Q9WUY7
28	82.5	13.4	225	4	Q96PU1
29	82.5	13.4	270	4	Q96PJ2
30	82.5	13.4	2115	12	Q86500
31	82	13.3	311	16	Q92RT2
32	82	13.3	386	4	Q9NQU0
33	82	13.3	392	4	Q75526
34	82	13.3	452	4	Q96SB6
35	82	13.3	802	5	Q9VJ74
36	81.5	13.2	163	12	Q9JG79
37	81.5	13.2	528	16	Q9RSJ1
38	81.5	13.2	2116	12	O40955
39	81.5	13.2	2301	10	Q9ATK5
40	81	13.1	148	4	Q96BZ7
41	81	13.1	394	2	Q9LO80
42	81	13.1	465	4	Q9BWD7
43	81	13.1	513	4	Q96H99
44	81	13.1	553	16	Q9RSN4
45	81	13.1	772	2	Q9AD03

ALIGNMENTS

RESULT 1

O60667 PRELIMINARY; PRT; 390 AA.

AC O60667; TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ANTI-FAS-INDUCED APOPTOSIS (REGULATOR OF FAS-INDUCED APOPTOSIS).

GN TOSO.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98246048; PubMed=9586636;

RA Hitoshi Y., Lorens J., Kitada S.I., Fisher J., Labarge M., Ring H.Z.,

RA Francke U., Reed J.C., Kinoshita S., Nolan G.P.;

RT "Toso, a cell surface, specific regulator of Fas-induced apoptosis in

RT T cells.";

RL Immunity 8:461-471(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=PRIMARY B-CELLS FROM TONSILS;

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF057557; AAC18830.1; -;

DR EMBL; BC006401; AAH06401.1; -;

DR InterPro; IPR003599; Ig.

DR InterPro; IPR003006; Ig_MHC.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00409; Ig; 1.

SQ SEQUENCE 390 AA; 43146 MW; FE91D217E9CA99C6 CRC64;

Query Match 100.0%; Score 616; DB 4; Length 390;
Best Local Similarity 100.0%; Pred. No. 7.3e-54;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRAVERKALSRARRLAVRMRALRESSQPRGRSQRPRSONNIYSACPRRARGADAAGTGE 60

```
Db 273 KRAVERKALSRARLAVRMRALESSQRPGRSPRPNNNIYSACPRRARGADAAGT 332
|||||
QY 61 AVPGCGAPLPAPLQVSESWLHAPSLKTSCEYSLYHQPAAMWEDSDDYINVPA 118
|||||
Db 333 AVPGCGAPLPAPLQVSESWLHAPSLKTSCEYSLYHQPAAMWEDSDDYINVPA 390
|||||

RESULT 2
Q9D8T1 ID Q9D8T1 PRELIMINARY; PRT; 422 AA.
AC Q9D8T1:
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE 1810037B05RIK PROTEIN.
GN 1810037B05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007714; BAB25207.1; -.
DR MGD; MGI:L916419; 1810037B05RIK.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00410; IG_like.1.
SQ SEQUENCE 422 AA; 47532 MW; 2597083A50AD8E6E CRC64;

Query Match 58.7%; Score 361.5; DB 11; Length 422;
Best Local Similarity 60.7%; Pred. No. 2.5e-28;
Matches 74; Conservative 12; Mismatches 31; Indels 5; Gaps 1;

QY 1 KRAVERKALSRARLAVRMRALESS-----QRPGRSPRPNNNIYSACPRRARGADA 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 284 KRAIQRSSRRAGELAMRRRGASRPFPPTQRDASQRPSPRPNNNIYSACPRRARGDS 343
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 56 AGTEAPVPGGAPLPAPLQVSESWLHAPSLKTSCEYSLYHQPAAMWEDSDDYIN 115
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 344 LGPAEAPLLNAPASASPQVLEAPWPHPTPSLKMSCETVSLGYPQAVNLEDPDSDYIN 403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 116 VP 117
:|
Db 404 IP 405
:|

RESULT 3
Q64561 ID Q64561 PRELIMINARY; PRT; 114 AA.
```

```
AC Q64561:
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE NUCLEAR TRANSITION PROTEIN 2 (TP-2).
GN TNP2 OR TP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HISTAR; TISSUE=TESTIS;
RX MEDLINE=97118932; PubMed=8954887;
RA Meitel A.R., Rao M.R.;
RT "Cloning of cDNA encoding rat spermatid protein TP2 and expression
RT in Escherichia coli.";
RL Protein Expr. Purif. 8:409-415(1996).
CC -!- FUNCTION: IN THE ELONGATING SPERMATIDS OF MAMMALS. THE CONVERSION
CC OF NUCLEOSOMAL CHROMATIN TO THE COMPACT, NONNUCLEOSOMAL FORM FOUND
CC IN THE SPERM NUCLEUS IS ASSOCIATED WITH THE APPEARANCE OF A SMALL
CC SET OF BASIC CHROMOSOMAL TRANSITION PROTEINS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN SPERMATID NUCLEAR
CC TRANSITION PROTEINS 2.
DR EMBL; U52958; AAB02693.1; -.
DR InterPro; IPR000678; TP2.
DR Pfam; PF01254; TP2; 1.
DR PROSITE; PS00970; TP2_1; 1.
DR PROSITE; PS00971; TP2_2; 1.
KW Chromosomal protein; DNA-binding; Metal-binding; Nuclear protein;
KW Nucleosome core; Spermatogenesis; Zinc.
SQ SEQUENCE 114 AA; 12848 MW; E9EFAF3BE3FC1A7C CRC64;

Query Match 15.6%; Score 96; DB 11; Length 114;
Best Local Similarity 32.5%; Pred. No. 0.023;
Matches 25; Conservative 10; Mismatches 36; Indels 6; Gaps 3;

QY 20 RMRALLESSQ-RPRGSPRPNNNIYSACPRRARGADAAG---TGEAPVPGGAPLPAPL 75
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 4 KMOSLPTTHPHSPHSRQSHNNQACSHHCRCQAGHSSSSSPSPGPTPKPTPM 63
|||||

QY 76 QVSESWLHAPSLKTSQ 92
|||:|||||
Db 64 HSRYSQ--SRPSHRGSC 78

RESULT 4
Q64390 ID Q64390 PRELIMINARY; PRT; 115 AA.
AC Q64390:
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE NUCLEAR TRANSITION PROTEIN 2 (TP-2).
GN TNP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96341725; PubMed=8720108;
RA Schluter G., Celik A., Obata R., Schlicker M., Hofferbert S.,
RA Schlung A., Adham I.M., Engel W.;
RT "Sequence analysis of the conserved protamine gene cluster shows that
RT it contains a fourth expressed gene.";
RL Mol. Reprod. Dev. 43:1-6(1996).
CC -!- FUNCTION: IN THE ELONGATING SPERMATIDS OF MAMMALS. THE CONVERSION
CC OF NUCLEOSOMAL CHROMATIN TO THE COMPACT, NONNUCLEOSOMAL FORM FOUND
CC IN THE SPERM NUCLEUS IS ASSOCIATED WITH THE APPEARANCE OF A SMALL
CC SET OF BASIC CHROMOSOMAL TRANSITION PROTEINS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
```

```
CC -!- SIMILARITY: STRONG, TO OTHER MAMMALIAN SPERMATID NUCLEAR
CC TRANSITION PROTEINS 2.
DR EMBL; 246939; CAA87064.1; -.
DR InterPro: IPR000678; TP2.
DR Pfam: PF01254; TP2; 1.
DR PROSITE; PS00970; TP2_1; 1.
DR PROSITE; PS00971; TP2_2; 1.
KW Chromosomal protein; DNA-binding; Metal-binding; Nuclear protein;
KW Nucleosome core; Spermatogenesis; Zinc.
SQ SEQUENCE 115 AA; 12948 MW; 29E180FF27894C1A CRC64;

Query Match 15.6%; Score 96; DB 11; Length 115;
Best Local Similarity 32.5%; Pred. No. 0.024; 36; Indels 6; Gaps 3;
Matches 25; Conservative 10; Mismatches 10;

QY 20 RMALESSO-RPRGSPRSONNIYSACPRRARGADAAG---TGEAPVPGGAPLPAPL 75
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 4 KMQLTTPHPHSSSRPOSHTNQACSHCHSCSQAGHPSSSSSPSPGPTKHPKPTM 63
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
QY 76 QVSESPWLHAPSLKTSK 92
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 64 HSRYP--SRPSHRGSC 78

RESULT 5
O46132 PRELIMINARY; PRT; 559 AA.
AC O46132;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NICOTINIC ACETYLCHOLINE RECEPTOR, ALPHA1 SUBUNIT (FRAGMENT).
GN NACHR.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridoidea; Acrididae; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP MEDLINE=98325051; PubMed=9660807;
RA Hermen B., Stetzer E., Thees R., Heiermann R., Schratzenholz A.,
RA Ebblinghaus U., Kretschmer A., Methfessel C., Reinhardt S.,
RA Maelicke A.;
RT "Neuronal nicotinic receptors in the locust locusta migratoria.";
RL J. Biol. Chem. 273:18394-18404(1998).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL; AJ000390; CAA04052.1; -.
DR InterPro: IPR000188; GABAA_receptor.
DR InterPro: IPR001175; Neur_chan.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEUOTR_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
FT TER
SQ SEQUENCE 559 AA; 62442 MW; 6EB9B33F2778B3DE CRC64;

Query Match 15.1%; Score 93; DB 5; Length 559;
Best Local Similarity 26.9%; Pred. No. 0.26;
Matches 36; Conservative 12; Mismatches 60; Indels 26; Gaps 3;

QY 1 KRAVERKALSRARRLAVRMRALES-----SORPRGSPRSONNIYSACPR 48
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 361 ERGLRRRGRGRGRAGAPGQLQRPANAHLOGRPLARQAPARRAAGQLPGARAAGR 420
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
QY 49 RARGADAAGT-----EAPVPGGAPLPAPLQVSESWLHA---PSLKTSCY 94
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 421 GAGGAAVGRGGVQEPATAAAATASGPGAVAPAGVRVSRPPAFPHSRCPPEVHRSCFC 480
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
QY 95 VSLYHQPAAMMEDS 108
```

```
Db 481 VRFTAEHTRMLEDS 494
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 6
O9WSV8 PRELIMINARY; PRT; 163 AA.
ID O9WSV8;
AC O9WSV8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DNA, COMPLETE GENOME, ISOLATE:TTV SANBAN.
OS "TT virus".
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TTV SANBAN;
RA Hijikata M.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TTV SANBAN;
RX MEDLINE=99335592; PubMed=10405352;
RA Hijikata M., Takahashi K., Mishiro S.;
RT "Complete circular DNA genome of a TT virus variant (isolate name
RT SANBAN) and 44 partial ORF2 sequences implicating a great degree of
RT diversity beyond genotypes.";
RL Virology 260:17-22(1999).
DR EMBL; AB025946; BAA82148.1; -.
DR InterPro: IPR004118; TT_ORF2.
DR Pfam; PF02957; TT_ORF2; 1.
SQ SEQUENCE 163 AA; 17179 MW; 794454FBD979B17E CRC64;

Query Match 14.5%; Score 89.5; DB 12; Length 163;
Best Local Similarity 31.1%; Pred. No. 0.15;
Matches 32; Conservative 5; Mismatches 39; Indels 27; Gaps 4;

QY 5 ERKALSRARRLAVRMRALESQRPRGSPRSONNIYSACPRRARGADAAG----- 57
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 21 KKKKLLSLRAPQAPRRAM--SWRPVHDAPCIERNWYACFRAHAGSCGCGNFIAHN 78
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58 -----TGEAPVFG---PGAP-----LPPAPLQVSESPW 82
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 79 LLAGRYGFTGGPPGPPGPPGTPQVQRASNSPAAPQQPPALPW 121
      :|::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 7
Q94H83 PRELIMINARY; PRT; 748 AA.
ID Q94H83;
AC Q94H83;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE HEAT SHOCK PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E.,
RA utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBA0090p23 genomic sequence.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084380; AAK52144.1; -.
KW Heat shock.
SQ SEQUENCE 748 AA; 83405 MW; 5CF14F897C71B7E2 CRC64;
```

```
Query Match 14.5%; Score 89.5; DB 10; Length 748;
Best Local Similarity 29.5%; Pred. No. 0.8;
Matches 31; Conservative 13; Mismatches 46; Indels 15; Gaps 4;

QY 7 KKAALRRRLAVRMRALESSQPRGSPRSONNIY-----SACPRARG-ADAAGTGE 60
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 122 KRALYDOKRLWLKR--NFSQTNKASAPGASNGFYNAANAASKVYTRGNKQKAGPAT 179
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 61 APVPGAPLPPAPLQVSESPWLHAPSLKTS-----EYVSLY 98
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 180 SSVRQRPPPPPPRQAPAPPAPPAKPTFTWTSCKNCKMNYELKVY 224
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 8
Q9RUB1 ID Q9RUB1 PRELIMINARY; PRT; 526 AA.
AC Q9RUB1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE ALGP-RELATED PROTEIN.
GN DRI1480
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Mofat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001992; AAF11045.1; -.
DR TIGR; DRI1480; -.
KW Complete proteome.
SQ SEQUENCE 526 AA; 55942 MW; A6C159D4612FE13F CRC64;

Query Match 14.4%; Score 88.5; DB 16; Length 526;
Best Local Similarity 30.2%; Pred. No. 0.69;
Matches 29; Conservative 14; Mismatches 34; Indels 19; Gaps 5;

QY 3 AYERKALRRRLAVRMRALESSQPRGSPRSONNIYACPRRARGADAAGTGE-- 60
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 386 AAESLHFSHRAQRI-IGTAWTESIEEPKATPAP-----AAPPAKQTQGTGTATW 436
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 61 APVPGAPLPPAPLQVSESPWLHAPSLK 89
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 437 GGSAPAPAPAAKGMESTSPAP-KPGDAPQSNAPVK 471
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 9
Q9NKP7 ID Q9NKP7 PRELIMINARY; PRT; 2936 AA.
AC Q9NKP7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE L712.2.
GN L712.2.
OS Leishmania major.
OC Eukaryota; Eulenzozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN;

RA Myler P.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC005930; AAF39727.1; -.
SQ SEQUENCE 2936 AA; 305212 MW; B5689E280242FE68 CRC64;

Query Match 14.2%; Score 87.5; DB 5; Length 2936;
Best Local Similarity 35.3%; Pred. No. 5.5;
Matches 30; Conservative 5; Mismatches 37; Indels 13; Gaps 3;

QY 12 RRARRLAVRMRALESSQPRGSPRSONNIYAC-----PRRARGADAAGTGEAPV 63
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 77 RKRELTAAATAIARGSPRGSAAAHAAAPLALLAPTRHPQRGRTATAAASKTTPS 136
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 64 PGGAPLPPAPLQVSESPWLHAPSL 88
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 137 PPKSAP-PPTPPPPSSP-----PSL 156
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 10
Q96JM3 ID Q96JM3 PRELIMINARY; PRT; 821 AA.
AC Q96JM3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE KIAA1802 PROTEIN (FRAGMENT).
GN KIAA1802.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
DR EMBL; AB058705; BAB47431.1; -.
FT NON TER 1
SQ SEQUENCE 821 AA; 90139 MW; 7948811E6E9F1DD8 CRC64;

Query Match 14.1%; Score 87; DB 4; Length 821;
Best Local Similarity 31.7%; Pred. No. 1.6;
Matches 33; Conservative 10; Mismatches 45; Indels 16; Gaps 4;

QY 9 ALSRRARRLAVRMRALESSQPRGSPRSONNIYACPRRARGADAAGTGEAPVPGPG 67
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 304 AVSPEPRPAPAVPGSWKPGPGSPRPMKSNPSASSGPMKPAKPAPSVS-----PGPW 357
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 68 APLEPAPLQVSESPWLHAPSL-----KTSCEYVSLYHQAPAMME 106
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 358 KPDP-----SVSPGPMKPTPSVSSASWKSWSVSPSSWKSPPASPE 397
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 11
Q30801 ID Q30801 PRELIMINARY; PRT; 166 AA.
AC Q30801;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE METHANOL OXIDATION GENES, MYAE, MYAH, MYAB, AND PMI-LIKE GENES
DE (FRAGMENT).
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1;
```

Q9VYTX7

AC Q9VYTX7 PRELIMINARY; PRT: 554 AA.

ID Q9VYTX7

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE EVG PROTEIN.

GN EVG OR CG10488.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI_Taxid=7227;

[1]

SEQUENCE FROM N.A.

RP STRAIN=BERKELEY.

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

RA Baillet R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Butchan M.R., Bouck J., Crokstein P., Brottier P.,

RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.G., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster K.A., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Spier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Shue E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svyrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -!- SIMILARITY: WITH OTHER HOMEOBOX PROTEINS.

DR EMBL: AE003541; AAF49918.1; -.

DR HSPSP; P06601; 1FJL.

DR FlyBase; FBgn0000625; eyg.

DR InterPro; IPR001356; Homeobox.

DR InterPro; IPR001523; Paired_box.

DR Pfam; PF00046; homeobox; 1.

DR Pfam; PF00292; PAX; 1.

DR SMART; SM00389; HOX; 1.

DR SMART; SM00351; PAX; 1.

DR PROSITE; PS00027; HOMEOBOX_1; 1.

DR PROSITE; PS50071; HOMEOBOX_2; 1.

KW DNA-binding; Homeobox; Nuclear protein.

SEQUENCE 554 AA; 59552 MW; 4BAA8D65FCEAE114 CRC64;

Query Match 13.8%; Score 85; DB 5; Length 554;

Best Local Similarity 25.2%; Pred. No. 1.6;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:28:11 : Search time 16.25 Seconds
(without alignments)
586.214 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLWPLYFLPVSGALRIL.....HQPAAMMEDSDDDYINVPA 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	99.6	390	3	US-08-961-564A-2
2	550	26.8	107	3	US-08-961-564A-4
3	198	9.6	771	3	US-08-434-000A-8
4	198	9.6	771	4	US-09-312-157-8
5	187	9.1	109	3	US-08-961-564A-9
6	185	9.0	769	3	US-08-434-000A-10
7	185	9.0	769	4	US-09-312-157-10
8	182	8.9	608	4	US-09-095-385-4
9	182	8.9	746	3	US-08-434-000A-4
10	182	8.9	746	4	US-09-312-157-4
11	182	8.9	757	3	US-08-434-000A-6
12	182	8.9	757	4	US-09-312-157-6
13	166.5	8.1	624	2	US-08-642-406A-22
14	166.5	8.1	624	4	US-09-199-534-22
15	166.5	8.1	773	3	US-08-434-000A-2
16	166.5	8.1	773	4	US-09-312-157-2
17	117.5	5.7	476	3	US-08-487-550-4
18	117	5.7	453	3	US-08-466-151-8
19	117	5.7	453	4	US-08-466-163B-8
20	113.5	5.5	2337	3	US-08-713-118-2
21	113.5	5.5	2337	4	US-09-452-007-2
22	112	5.5	451	2	US-08-887-352B-14
23	112	5.5	451	2	US-08-887-352B-16
24	112	5.5	451	3	US-08-466-151-65
25	112	5.5	451	4	US-09-109-207C-14
26	112	5.5	451	4	US-09-109-207C-16
27	112	5.5	451	4	US-09-296-005-14

28	112	5.5	451	4	US-09-296-005-16	Sequence 16, Appl
29	111	5.4	476	2	US-08-378-939-10	Sequence 10, Appl
30	109.5	5.3	2339	1	US-08-455-543A-47	Sequence 47, Appl
31	109.5	5.3	2339	4	US-08-223-305C-47	Sequence 47, Appl
32	109.5	5.3	2339	4	US-09-268-163-6	Sequence 6, Appl
33	109.5	5.3	2343	4	US-09-268-163-4	Sequence 4, Appl
34	107	5.2	292	4	US-09-345-468-18	Sequence 18, Appl
35	107	5.2	313	4	US-09-345-468-16	Sequence 16, Appl
36	107	5.2	449	1	US-08-458-516-13	Sequence 13, Appl
37	106	5.2	390	2	US-08-887-352B-18	Sequence 18, Appl
38	105	5.1	451	2	US-09-109-207C-18	Sequence 18, Appl
39	105	5.1	451	4	US-09-282-505-2	Sequence 2, Appl
40	105	5.1	451	4	US-09-054-255-2	Sequence 2, Appl
41	105	5.1	451	4	US-09-296-005-18	Sequence 18, Appl
42	105	5.1	451	4	US-08-347-254-1	Sequence 1, Appl
43	104	5.1	534	1	US-08-464-463-1	Sequence 1, Appl
44	104	5.1	534	2	US-08-157-101A-7	Sequence 7, Appl
45	103.5	5.0	459	1		

ALIGNMENTS

RESULT 1
US-08-961-564A-2
; Sequence 2, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: PICRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-2

Query Match 99.6%; Score 2047; DB 3; Length 390;
Best Local Similarity 99.7%; Pred. No. 2e-178;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLWPLYFLPVSGALRLILPEVKVEGELGSGSVTIKCPLEPMHVRIYLCREMAGSGTCGT 60
DB 1 MDRWLWPLYFLPVSGALRLILPEVKVEGELGSGSVTIKCPLEPMHVRIYLCREMAGSGTCGT 60
QY 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQLTSDSGSVYACGAGMNTDRGKTQKVTL 120
DB 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQLTSDSGSVYACGAGMNTDRGKTQKVTL 120
QY 121 NVHSEYPSWEQPMPTPKWFHLYLQMPAYASSKKEVTRVTPAQRGKVPVPHHSSP 180
DB 121 NVHSEYPSWEQPMPTPKWFHLYLQMPAYASSKKEVTRVTPAQRGKVPVPHHSSP 180
QY 181 TTQITHRPRVSRASSVAGDKPRTFLPSTASKISALEGLKFPQTPSYNHHTRLHQRALD 240
DB 181 TTQITHRPRVSRASSVAGDKPRTFLPSTASKISALEGLKFPQTPSYNHHTRLHQRALD 240
QY 241 YGQSQRGQGFHILPTILGLFILALLGLVWRAVERKALSRRLAVRMRALESSQ 300
DB 241 YGQSQRGQGFHILPTILGLFILALLGLVWRAVERKALSRRLAVRMRALESSQ 300
QY 301 RPRGSPRPSQNTYSACPRRARGADAAGTGEAPVPGCAPLPAPLQVSESPWLHAPSL 360
DB 301 RPRGSPRPSQNTYSACPRRARGADAAGTGEAPVPGCAPLPAPLQVSESPWLHAPSL 360
QY 361 KTSCEVSVLYHOPAAHMDSDSDYINVPA 390
DB 361 KTSCEVSVLYHOPAAHMDSDSDYINVPA 390

RESULT 2

US-08-961-564A-4
; Sequence 4, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUENH, ALEMSED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F.
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-564A-4
Query Match 26.8%; Score 550; DB 3; Length 107;
Best Local Similarity 99.0%; Pred. No. 6.6e-43;
Matches 104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDRWLWPLYFLPVSGALRLILPEVKVEGELGSGSVTIKCPLEPMHVRIYLCREMAGSGTCGT 60
DB 1 MDRWLWPLYFLPVSGALRLILPEVKVEGELGSGSVTIKCPLEPMHVRIYLCREMAGSGTCGT 60
QY 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQLTSDSGSVYACG 105
DB 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQLTSDSGSVYACG 105
RESULT 3
US-08-434-000A-8
; Sequence 8, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; TOPOLOGY: DESCRIPTION: Mouse Polyimmunoglobulin Receptor
US-08-434-000A-8
Query Match 9.6%; Score 198; DB 3; Length 771;
Best Local Similarity 31.5%; Pred. No. 1.2e-09;
Matches 53; Conservative 23; Mismatches 68; Indels 24; Gaps 5;
QY 10 FLVSGALRLILPEVKVEGELGSGSVTIKCPLE----MHVRIYLCREMAGSGTCGTVVSTT 65

Db 13 FSGVSTKSPFGQVSSIEGDSVITCYYPDTSVNRHTRKYWCQGA-SGMCITLISN 71
QY 66 NFIKAEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLVHSE 125
Db 72 GYLSKEYSGRANLINFENNTFVINEQLTQDDTGSYKCGLG-TSNRGLSFDVSLEV--- 127
QY 126 YEPSWEEQMPETPKWFHLPYLFQMPAYASSSKFVTRVTTPAORGVKVP 173
Db 128 -----SQVPELPSDTHV-----YTKDIGRNVITIECPFKRENVP 160

RESULT 4
US-09-312-157-8
; Sequence 8, Application US/09312157
; Patent No. 6303341
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,157
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,000
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-351
; SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: Mouse Polyimmunoglobulin Receptor

US-09-312-157-8
Query Match 9.6%; Score 198; DB 4; Length 771;
Best Local Similarity 31.5%; Pred. No. 1.2e-09;
Matches 53; Conservative 23; Mismatches 68; Indels 24; Gaps 5;
QY 10 FLPVSGALRLLPEVKEGELGSGVTIKCPLE---MHVRIYLCREMAGSGTCGTGVVSTT 65
Db 13 FSGVSTKSPFGQVSSIEGDSVITCYYPDTSVNRHTRKYWCQGA-SGMCITLISN 71
QY 66 NFIKAEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLVHSE 125
Db 72 GYLSKEYSGRANLINFENNTFVINEQLTQDDTGSYKCGLG-TSNRGLSFDVSLEV--- 127

QY 126 YEPSWEEQMPETPKWFHLPYLFQMPAYASSSKFVTRVTTPAORGVKVP 173
Db 128 -----SQVPELPSDTHV-----YTKDIGRNVITIECPFKRENVP 160
RESULT 5
US-08-961-564A-9
; Sequence 9, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUIJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 23-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-9

Query Match 9.1%; Score 187; DB 3; Length 109;
Best Local Similarity 38.3%; Pred. No. 7.7e-10;
Matches 41; Conservative 19; Mismatches 41; Indels 6; Gaps 3;
QY 21 PEVKVEGELGSGVTIKCPLE---EMHVRIYLCREM-AGSGTCGTGVVSTTNTFIKAEYKGR 75
Db 4 PE-EVNSVEGNSVITCYYPDTSVNRHTRKYWCQRCARGGLCITLISSEGYVSSKYAGR 62
QY 76 VTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLV 122
Db 63 ANLTNPFENGTFVNVIAQLSQDDSGRYKCGLGINSRLGLSFDVSLEV 109
RESULT 6
US-08-434-000A-10
; Sequence 10, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER

CLASSIFICATION:
PRIOR APPLICATION DATA: 60/050,969
FILING DATE: 19-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Canady, Karen S
REGISTRATION NUMBER: 39,927
REFERENCE/DOCKET NUMBER: 30435.45USU1
TELEPHONE: 310 445-1140
TELEFAX: 310 445-9031
TELEX:

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-095-385-4

Query Match 8.9%; Score 182; DB 4; Length 608;
Best Local Similarity 37.4%; Pred. No. 2.4e-08;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PVSGALRILPEVKVEGLGSGVTIKCPLP-----EMHVRIYLCREMAGSGTCGTVVSTTNF 67
Db 21 PIFG-----PE-EVNSVEGNSVITCYPPTSVNRHTRKYWCROGARGG-CITLISSEGY 73
QY 68 IKAERYKRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLV 122
Db 74 VSSKYAGRANLTNPENGTFVFNIAQLSQDSDSGRYKCGLGINS-RGLSPDVSLEY 127

RESULT 9

US-08-434-000A-4
Sequence 4, Application US/08434000A
Patent No. 6046037

GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
APPLICANT: K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,000A
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 1
APPLICATION NUMBER: 08/367,395
FILING DATE: 12/30/94
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400

TELEFAX: (619) 552-0159
TELEX: 67-3510
TELEX: SEQUENCE LISTING
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
TOPOLOGY: DESCRIPTION: Human Polyimmunoglobulin Receptor
US-08-434-000A-4

Query Match 8.9%; Score 182; DB 3; Length 746;
Best Local Similarity 37.4%; Pred. No. 3.2e-08;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PVSGALRILPEVKVEGLGSGVTIKCPLP-----EMHVRIYLCREMAGSGTCGTVVSTTNF 67
Db 3 PIFG-----PE-EVNSVEGNSVITCYPPTSVNRHTRKYWCROGARGG-CITLISSEGY 55
QY 68 IKAERYKRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLV 122
Db 56 VSSKYAGRANLTNPENGTFVFNIAQLSQDSDSGRYKCGLGINS-RGLSPDVSLEY 109

RESULT 10

US-09-312-157-4
Sequence 4, Application US/09312157
Patent No. 6303341

GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
APPLICANT: K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,157
FILING DATE: 14-May-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/434,000
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-351

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: Human Polyimmunoglobulin Receptor
SEQU ID NO: 4

US-09-312-157-4

Query Match 8.9%; Score 182; DB 4; Length 746;
Best Local Similarity 37.4%; Pred. No. 3.2e-08;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;
QY 12 PVSGALRILPEVKVEGELGGVTIKCLP-----EMHVRIYLCREMAGSGTGTWSTNF 67
Db 3 PIFG-----PE-EYNSVEGNSVITCYPTTSVNRHTRKYWCROGARGG-CITLISSEGY 55
QY 68 IKAETKGRVTLKQYPRKNFLVEVTQLTESDSGVYACGAGMTDRGKTQKVTLVN 122
Db 56 VSSKYAGRANLTNPENGTEVYVNIQLSDSDSGRYKCGLGINS-RGLSFDVSLEV 109

RESULT 11

US-08-434-000A-6
; Sequence 6, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE: 12/30/94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application 1
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 757 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-434-000A-6

Query Match 8.9%; Score 182; DB 3; Length 757;
Best Local Similarity 34.1%; Pred. No. 3.2e-08;
Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;
QY 17 LRILPEVKVEGEL-----GGSVTIKCLP-----EMHVRIYLCREMAGSGTGTWV 62
Db 10 LAIFPVVSMKSPIFGPEEVSVEGSRVSIKYYPTTSVNRHTRKYWCROGA-QGRCTTLI 68

QY 63 STTNFIKAETKGRVTLKQYPRKNFLVEVTQLTESDSGVYACGAGMTDRGKTQKVTLVN 122
Db 69 SSEGYSDDYVGRANLTNPESGTFVVDISHLTHKDSGRYKCGLISS-RGLNFDVSLEV 127
QY 123 HSE 125
Db 128 SQD 130

RESULT 12

US-09-312-157-6
; Sequence 6, Application US/09312157
; Patent No. 6303341
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,157
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,000
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-351
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 757 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Bovine Polyimmunoglobulin Receptor
; US-09-312-157-6

Query Match 8.9%; Score 182; DB 4; Length 757;
Best Local Similarity 34.1%; Pred. No. 3.2e-08;
Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;
QY 17 LRILPEVKVEGEL-----GGSVTIKCLP-----EMHVRIYLCREMAGSGTGTWV 62
Db 10 LAIFPVVSMKSPIFGPEEVSVEGSRVSIKYYPTTSVNRHTRKYWCROGA-QGRCTTLI 68
QY 63 STTNFIKAETKGRVTLKQYPRKNFLVEVTQLTESDSGVYACGAGMTDRGKTQKVTLVN 122
Db 69 SSEGYSDDYVGRANLTNPESGTFVVDISHLTHKDSGRYKCGLISS-RGLNFDVSLEV 127
QY 123 HSE 125


```
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 1
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; TELETYPE: SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 773 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Rabbit polyimmunoglobulin receptor
US-08-434-000A-2

Query Match      8.18; Score 166.5; DB 3; Length 773;
Best Local Similarity 42.7%; Pred. No. 8.6e-07;
Matches 44; Conservative 11; Mismatches 41; Indels 7; Gaps 4;

QY 30 GGSVTIKCPLP---EMHVIYLCREMAGSGTGTWVSTTNFIKAHYKGRVTLKQYPRKN 85
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 GDSVSICTCYPTTSVTRHSRKFWCRE-EESGRCVTLAS-TGYTSQEYSGRGKLTDFPDKG 96
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 86 LFLVETVLTEDSGVYACGAGMNTDRGKTQKVTLNHSEYEP 128
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 97 EFVTVDTQDTSQSYKCGVGN-GRGLDFGVNVLVSQKPEP 138
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: October 28, 2002, 17:33:10
Job time : 19.25 secs
```



```

QY 222 POTPSYNNHTRL---HRQALDYGSGREGOG----- 251
Db 438 FO-PSQAQSCSSDHSAGSAANTDSGVMTGNAGGSTDLSLDEGPCIAGTGLSLR 496
QY 252 -----PHILIPT-ILGLFL--LALLGLVVKRAVERRKALSR-----ARRLAVR 292
Db 497 RLENLRDLQHLRIGTRGLKPLSLSHGTGTVSRVGPVPPVSLERRSSSSSISSAYTVS 556
QY 293 WRALESSORPGSPRPRSONNIYSACPR-----RARGADAAGTGEAPVPGP-----GAP 341
Db 557 RSSLASPPPPSPENGAGSSLLPQAQHYLLRARYASARGGGTPTAASLDRIGCLP 616
QY 342 LPP-----APLQVSESPWLHAPSLKTSCEVVSLYHOPAMMEDS 380
Db 617 MPPWRSAEYGVNPNAGVTRRASDPAQADR-----APARVOREKSLGCVHTPTTVAGG 673
QY 381 DS-DDYI 386
Db 674 QNFDPYL 680

RESULT 7
EDBEGA
immediate-early protein IE175 - Marek's disease virus (strain GA)
N:Alternate names: infected cell protein 4
C:Species: Marek's disease virus
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999
C:Accession: A42991
R:Anderson, A.S.; Francesconi, A.; Morgan, R.W.
Virology 189, 657-667, 1992
A:Title: Complete nucleotide sequence of the Marek's disease virus ICP4 gene.
A:Reference number: A42991; MUID:92351564
A:Accession: A42991
A:Molecule type: DNA
A:Residues: 1-1415 <AND>
A:Cross-references: GB:M75729; NID:g330950; PIDN:AAA46111.1; PID:g330951
C:Superfamily: herpesvirus immediate-early protein IE175
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 5.5%; Score 113.5; DB 1; Length 1415;
Best Local Similarity 21.1%; Pred. No. 3.3;
Matches 70; Conservative 37; Mismatches 108; Indels 117; Gaps 14;

QY 104 CGAGMNTDRGKTQKVTLVNHSEYEPS--WEEQMPETPKWFHLPYLFQMPAYASSSKFVT 161
Db 44 CNAAMDDD-QQMPFPLPVPDWNPNSTDWRPSPRSGPK--KDFCGDLPAFLTSG--P 96
QY 162 RVTPPAORGKVPVHSSPTTQITHRP-----VS 191
Db 97 RLTFPSS-GRMSELPHTTSSPRSPRGPTSPSNEHIIISPPRNPFSNTHRNVGHS 155
QY 192 RASVAGDKPRFLPSTT-----ASKISALEGLLKPTQPSYNNHTRLHRORALDYGSGG 246
Db 156 RSPSSSSSSSSSSSSSLVLSPPSSSRSPSPRPRADSSSRPRGRGSRGRSG 215
QY 247 REGOGHILPTILGLFLALLGLLVKRAVERRKALSRARRLAVMRALESSORPGSP 306
Db 216 PQSG-----RKA-SPTRKLEDEYLPQETANRRGG 247
QY 307 RPARSONNIYSACPRRGA-----DAAGT--GEAPVPGCAPLP 345
Db 248 RPRGRP-----PKSGRAVORNDIQVTSSGLADTSYDLCGSVWVEVPLPPG----- 295
QY 346 PLOVSESPWL-----HAPSLKTSCEVVSLYHQ 372
Db 296 -----RCWFGGLGHRQALTDSPETVEAIHR 321

RESULT 8
CMRF-35 antigen - human
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
```

```

C:Accession: I37243
R:Jackson, D.G.; Hart, D.N.; Starling, G.; Bell, J.I.
Eur. J. Immunol. 22, 1157-1163, 1992
A:Title: Molecular cloning of a novel member of the immunoglobulin gene superfamily h
A:Reference number: I37243; MUID:92249405
A:Accession: I37243
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-224 <RES>
A:Cross-references: EMBL:X66171; NID:g396169; PIDN:CAA46948.1; PID:g396170
C:Genetics:
A:Gene: CMRF35
```

```

Query Match 5.5%; Score 112.5; DB 2; Length 224;
Best Local Similarity 22.5%; Pred. No. 0.49;
Matches 64; Conservative 33; Mismatches 100; Indels 87; Gaps 12;
```

```

QY 8 LYFLPVSGALRILPEVKVEGELGSGVTIKCPLPEMH--VRIYLCREMAGSGTCGTGVYSTT 65
Db 14 LLLLVPGYFPLSHPMVTAVGPGVSGSLVQCYEKEHRTLKFWCRP-PQILRCDKIVETK 72
QY 56 NFIKAERYKRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLVNHSE 125
Db 73 G-SAGRRGRVSIKSDSPANLSFTVTLENLTEDAGTYWCGV----- 112
QY 126 YEPSWEEQMPETPKW---PHLPYLFQMPAYASSSKFVTRVTPPAORGKVPVHHSSPTT 182
Db 113 -----DTP-WLRDFHDPV-----EVEVSVPFAGTTTA-----SSP-- 142
QY 183 QITHRPVSRASSVAGDKPTELPSTASKISALEGLLKQTPSPSYNNHTRLHRORALDYG 242
Db 143 -----QSSMGTSGPPTKLPVHTWPSVT-----RKDSEPSPH----- 174
QY 243 SOSGREGQGFILPTILGL-FLALLGLLVKRAVERRKALSR 285
Db 175 --PGSLFSNVRFLLVLELPLLSMLGAVLWNRQPSRSRQ 216
```

RESULT 9

```

T37056
probable multi-domain beta keto-acyl synthase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: T37056
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: 221620
A:Accession: T37056
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2082 <SEE>
A:Cross-references: EMBL:AL109747; PIDN:CAB52353.1; GSPDB:GN00070; SCODEB:SCJ21.07
A:Experimental source: strain A3(2)
C:Genetics:
```

```

A:Gene: SCODEB:SCJ21.07
C:Superfamily: Streptomyces coelicolor probable polyketide synthase; 3-oxoacyl-[acyl-
acyl-carrier-protein] S-malonyltransferase homology
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
P:48-479/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
P:1229/Binding site: phosphopantetheine (Ser) (covalent) #status predicted
```

```

Query Match 5.5%; Score 112; DB 2; Length 2082;
Best Local Similarity 24.1%; Pred. No. 6.6;
Matches 73; Conservative 20; Mismatches 124; Indels 86; Gaps 12;
```

```

QY 105 GAGMNTDRGKTQKVTLV-NVHSEYEPSWEEQMPETPKWFHLPYLFQMPAYASSSKFVTRV 163
Db 917 GAGATFHAALGELAVRGVPLDNLTPYASPVTFAKEREPRM-----TVK 963
QY 164 TTPAORGKVPV--VHHSSPTTQITHRPVSRASSVAGDKPTELPSTTASIKSLEGLLK 221
Db 964 ISGANYGQLPPSGPHTVETPAPDHGPRPDVAPVAAPVAVALPSQAGDV----- 1015
```



```
Query Match          5.4%; Score 111; DB 2; Length 801;
Best Local Similarity 20.6%; Pred. No. 2.7;
Matches 72; Conservative 36; Mismatches 157; Indels 84; Gaps 12;

QY 30 GGSVITKPLPEMHVRVILCLREMGSGTCGTGVYSTNFKAEYKGRVTLKQYPRKNLFLV 89
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 315 GDAATAPASSEADAAATDSSAASADTTAALVDTS---SEHAAEST---EAPATDIAAT 369

QY 90 EVTLQTESDSGVYACGAGMNTDRGKTQKVLNVHSEYEPSWEEQMP-----ETPKWFH 143
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 370 ETPAPSPVAPVADAAGYDSSPSIETP-----APAAEDTPAPASAAABETPA--P 420

QY 144 LPYLQMPAYSSKFEVTRVTPAQGRKVPVHVHSSPTTQITHPRVSRASSVAGDKPRT 203
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 APAEETPAPETAS-----AAPDAAGAAPADVAAPADVATTAPETSSAQAAGSYD-- 472

QY 204 FLPTSTASKISALEGLLKQTPSYNHHTRLHRQALDYGSOGREGOGFHILPTILGLF 263
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 473 -VPSEFASEVTA-----PIVESATEAPSDSAAPTG-PAASEPAPAPIEAP----- 515

QY 264 LLALLGLVYKRAVERRRKALSRRLAVRMRALESSQRPGSRPRSQ-----NNIY 315
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 516 -----ATDAATLETAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 549

QY 316 SACPFRAGAD-----AAGTGEAPVPGAPLPAPAPLOVSESPWLHAPS 359
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 550 EETPAPAPAADETAPAPAPAAETPAPAPAAETPAPAPAAETPAPATAPA 598

RESULT 13
A42566
omega-conotoxin-sensitive N-type calcium channel alpha 1B-1 subunit (alternatively spliced)
C:Species: Homo sapiens (man)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1998
C:Accession: A42566
R:Williams, M.E.; Brust, P.F.; Feldman, D.H.; Patthi, S.; Simerson, S.; Maroufi, A.; McCoy
Science 257, 389-395, 1992
A:Title: Structure and functional expression of an omega-conotoxin-sensitive human N-type
A:Reference number: A42566; MUID:92335886
A:Accession: A42566
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-2339 <WIL>
A:Experimental source: IMR32, hippocampus
A>Note: sequence extracted from NCBI backbone (NCBIP:109168)
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match          5.4%; Score 110.5; DB 2; Length 2339;
Best Local Similarity 20.0%; Pred. No. 9.7;
Matches 82; Conservative 49; Mismatches 139; Indels 139; Gaps 18;

QY 75 RVTLKQYPRKNLFLVEVTLQTESDSGVYACGAGMNTDRGKTQKVLNVHSEYEPSWEEQ 134
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1894 KATLEOTPAVLARGAVFLRQKSSSTLSNGGATQNGESGKESV-----SWQTOR 1943

QY 135 MPETPKWHLPL-----YLFQMPAYASS-----KFVTR-----VTTTP 166
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1944 TDAPHEAPPLERGHSTETIPVGRSGALVDVQMSITRRQDPQEPQGLSQORASWP 2003

QY 167 AQRGKVPVPHHSSP-----TTQTHPRVSRASSVAGDKPRTFLPTSTASKISALEGLLK 222
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2004 RLAAETQPVTDASPMKRSTSLAQPRGTHLCSTTPDRP-----PPSQAS----- 2048

QY 223 QTPSYNHHTRLHR-----QRALDYGSOGREGOGFHILPTILGLFLLALLG----- 269
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2049 ---SHHHHRRCHRRDRKORSLEKGPSLSADMDG---APS-----SAVGPLPPCEQ 2094

QY 270 -LVVKAVERKALSRRLAVRMRALESSORP-----RGSPPRSONNYSAC 318
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2095 PTOCRREYRQERGRQER---RQPSSESSEKQRYCDRFGREPKPKPSLSHTPS 2151

QY 319 PRRAR-----GADAAGTGEAPVPG-----PGAPLPAPAPAPAPAPAPAPAPAPAP 355
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 2152 PTAQBPQPHPGSGSVNGSPLLSTSGASTPGQRORQLPQPTLTPRPSTIYKTTANSSPI 2211
QY 356 HAPSLKTS-----CEYVSLYH-----QPAAMMEDSDDDYI 386
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2212 HFAQAQTSLPAPFSPGRSLSRGLSEHNALLQRPDLSQPLAPGSRIGSDPYL 2260

RESULT 14
B75310
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: B75310
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: B75310
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-528 <WHIT>
A:Cross-references: GB:AE002048; GB:AE000513; MID:g6459929; PIDN:AAF11681.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2133
A:Map position: 1

Query Match          5.2%; Score 107.5; DB 2; Length 528;
Best Local Similarity 25.7%; Pred. No. 3;
Matches 61; Conservative 25; Mismatches 96; Indels 55; Gaps 11;

QY 149 QMPAYASS-----KFVTRVTPAQGRKVPVHVHSSPTTQITHPRVSRASSVAGDKPRT 203
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 324 QAQARHPSSARRPCAFRCRSKPRGRQRQPV--TRPSSHVTRRRRPATRPSPGRRPST 381

QY 204 ----FLPTSTASKISALEGLLKQTPSYNHHTRLHRQALDYGSOGREGOGFHILPTI 259
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 392 PVTGWMPASATGCRUSA-----PRCRRTATRCRTATGCGRTSPRCGPSSGSC---- 428

QY 260 LGLFLALLGLLVYKRAVER-----RKALSRRLAVRMRALES-SQRPGSRPRRSO 311
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 429 -----RAATRRSPRTSPRRARASRSTRPTIPAPAANSASAAPNSPTRKTN 474

QY 312 NNIYSACPRRRAGADAGTGEAPVP--GPGAPLPAPLOVSESPW--LHAPSLKTS 364
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 475 WSTPGWCPRSAASTPSSRSFGAPPVNGFG-PEPTA-----RRRGWGRHAPRMKPN 526

RESULT 15
T02345
hypothetical protein KIAA0324 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 05-Nov-1999
C:Accession: T02345
R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
White, S.; Ueng, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
submitted to the EMBL Data Library, March 1998
A:Description: Sequencing of human chromosome 16p13.3.
A:Reference number: Z14664
A:Accession: T02345
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1791 <RIC>
A:Cross-references: EMBL:AC004493; MID:g2996648; PIDN:AAC08453.1; PID:g2996650
C:Genetics:
A:Map position: 16
A:Introns: 1610/2; 1706/2
A>Note: KIAA0324

Query Match          5.2%; Score 107; DB 2; Length 1791;
Best Local Similarity 23.5%; Pred. No. 13;
```

```
Matches 56; Conservative 30; Mismatches 100; Indels 52; Gaps 8;
QY 114 KTKVTLNVHSEYEPSEQMPETPKWFLPYLFQMPAYASSSKXEVTVTRVTPAORGKVP 173
Db 696 KSRARGSRSSPEPKTKRTPPRRSSRSPELTRKARLSRRSRAS--SSPETRSTRTP 753
QY 174 PVHSSPTTQITHRPRVSRAS--SVAGDKPRTFLPSTTA-----SKISALEGLLKPTPSY 227
Db 754 PRHRSFSVSSPEPAEKSRSSRRRSASSPT--KTTSSRGRSPSPKPRGLQRSRSR 810
QY 228 NHTRLRHRQALDYGQS-----GREGQGFHILPTILGLFLALLGL 270
Db 811 REKTRTTRRRDRSGSQSTSRRRQRSRSRVTRRRRGSGYH----- 853
QY 271 VVKRAVERRKALSRARRLAVRMALSSQ--RPRGSPRP--RSONNIYSACPRAR 323
Db 854 --SRSPARQESRTSSRRRGRSPTPTSKKRSRSTSPAPWKRKRSPATHRSR 909
```

Search completed: October 28, 2002, 17:33:59
Job time : 26.4435 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 11.0081 seconds
(without alignments)
1371.778 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLPLYLPLVSGALRIL.....HQPAAMMEDSDSDYINVPA 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	195	9.5	771	1 PIGR_MOUSE	O70570 mus musculu
2	185	9.0	769	1 PIGR_RAT	P15083 rattus norv
3	182	8.9	757	1 PIGR_BOVIN	P81265 bos taurus
4	182	8.9	764	1 PIGR_HUMAN	P01833 homo sapien
5	166.5	8.1	773	1 PIGR_RABIT	P01832 oryctolagus
6	121	5.9	102	1 PIGR_PIG	Q23244 sus scrofa
7	117	5.7	1106	1 GLI1_HUMAN	P08151 homo sapien
8	113.5	5.5	1415	1 ICP4_HSVMG	Q02362 marek's dis
9	112.5	5.5	224	1 CM35_HUMAN	Q08708 homo sapien
10	111.5	5.4	390	1 VGLI_HSV11	P06487 herpes simp
11	109.5	5.3	2339	1 CCAB_HUMAN	Q00975 homo sapien
12	102	5.0	530	1 PVR2_MOUSE	P32507 mus musculu
13	101.5	4.9	3530	1 MY15_HUMAN	Q90kn7 homo sapien
14	101	4.9	2339	1 CCAB_RABIT	Q05152 oryctolagus
15	100.5	4.9	876	1 AREA_EMENI	P17429 emericeella
16	100.5	4.9	1616	1 APXL_HUMAN	Q13796 homo sapien
17	100	4.9	384	1 DUS5_HUMAN	Q16690 homo sapien
18	99	4.8	430	1 CD19_CAVPO	P25917 cavia porce
19	98.5	4.8	116	1 STP2_RAT	P11101 rattus norv
20	98.5	4.8	341	1 FCQ2_CAVPO	Q60513 cavia porce
21	98.5	4.8	847	1 CD22_HUMAN	P20273 homo sapien
22	98	4.8	554	1 CSF1_HUMAN	P09603 homo sapien
23	97	4.7	3649	1 ACVS_NOCLA	P27743 nocardia la
24	95.5	4.6	538	1 PVR2_HUMAN	Q92692 homo sapien
25	95.5	4.6	1833	1 ZEP2_HUMAN	P31629 homo sapien
26	95	4.6	862	1 SM4D_HUMAN	Q92854 homo sapien
27	94.5	4.6	275	1 IAP_GVCP	P41436 cydia pomon
28	94.5	4.6	3149	1 TEGU_EBV	P03186 epstein-bar
29	94	4.6	825	1 IL4R_HUMAN	P24394 homo sapien
30	94	4.6	1183	1 DRPL_RAT	P54258 rattus norv
31	94	4.6	1490	1 CRK7_HUMAN	Q90ny4 homo sapien
32	94	4.6	2440	1 NCRL_HUMAN	O75376 homo sapien
33	93.5	4.5	346	1 ST50_YEAST	P25344 saccharomyc

RESULT 1																	
PIGR_MOUSE	34	93	4.5	355	1	PLK_CHICK	P07354	gallus gall									
ID	PIGR_MOUSE	35	93	387	1	PIGM_RHOSO	P26698	rhodococcus									
AC	O70570;	36	92.5	4.5	692	1	RFX2_MOUSE	P48379	mus musculu								
DT	30-MAY-2000 (Rel. 39, Created)	37	92.5	4.5	1174	1	KPCI_COCHE	O42632	cochliobolu								
DT	01-MAR-2000 (Rel. 39, Last sequence update)	38	92.5	4.5	1906	1	KMLS_CHICK	P11799	gallus gall								
DT	01-MAR-2002 (Rel. 41, Last annotation update)	39	92	4.5	283	1	EXTN_SORBI	P24152	sorghum bic								
DE	PolymERIC-IMMUNOGLOBULIN RECEPTOR PRECURSOR (Poly-IG RECEPTOR) (PIGR)	40	92	4.5	475	1	TFEB_MOUSE	Q9r210	mus musculu								
DE	[Contains: Secretory component].	41	92	4.5	1638	1	BRM_DROME	P25439	drosophila								
GN	PIGR.	42	92	4.5	2150	1	SDC3_CAEEL	P34706	caenorhabdi								
OS	Mus musculus (Mouse).	43	91.5	4.5	350	1	VGLI_PVRRI	P07646	pseudorhabdi								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	44	91.5	4.5	534	1	CNRD_CAEEL	P41830	caenorhabdi								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	45	91.5	4.5	872	1	FPL_MYTCO	Q25434	mytilus cor								
OX	NCBI_TaxID=10090;																
RN	[1]																
RP	SEQUENCE FROM N.A.																
RC	STRAIN=C57BL/6 X CBA; TISSUE=Liver;																
PX	MEDLINE=95138517; PubMed=7836758;																
RA	Piskurich J.F., Blanchard M.H., Youngman K.R., France J.A.,																
RA	Kaetzel C.S.;																
RT	"Molecular cloning of the mouse polymeric Ig receptor. Functional																
RT	regions of the molecule are conserved among five mammalian species.";																
RL	J. Immunol. 154:1735-1747(1995).																
RN	[2]																
RP	SEQUENCE FROM N.A.																
RC	STRAIN=129/SVJ;																
PX	MEDLINE=98072444; PubMed=9409786;																
RA	Martin M.G., Gutierrez E.M., Lam J.T., Li T.W.H., Wang J.;																
RA	"Genomic cloning and structural analysis of the murine polymeric																
RT	receptor (pigr) gene and promoter region.";																
RL	Gene 201:189-197(1997).																
RN	[3]																
RP	SEQUENCE FROM N.A.																
RC	STRAIN=129; TISSUE=Liver;																
RT	de Groot N., Vollebregt E., Lee S.H., Verbeet M.P., de Boer H.A.;																
RA	"Molecular cloning and exon-intron organization of the gene encoding																
RT	the murine polymeric immunoglobulin receptor.";																
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.																
CC	-!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE																
CC	BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN																
CC	TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.																
CC	DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE																
CC	EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE																
CC	TRANSMEMBRANE SEGMENT (BY SIMILARITY).																
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.																
CC	-!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.																
CC	-!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.	-----															
This SWISS-PROT entry is copyright. It is produced through a collaboration																	
between the Swiss Institute of Bioinformatics and the EMBL outstation -																	
the European Bioinformatics Institute. There are no restrictions on its																	
use by non-profit institutions as long as its content is in no way																	
modified and this statement is not removed. Usage by and for commercial																	
entities requires a license agreement (See http://www.isb-sib.ch/announce/																	
or send an email to license@isb-sib.ch).																	

ALIGNMENTS

```
DR EMBL; U06431; AAA67440.1; -.
DR EMBL; U83434; AAC53585.1; -.
DR EMBL; U83427; AAC53585.1; JOINED.
DR EMBL; U83428; AAC53585.1; JOINED.
DR EMBL; U83429; AAC53585.1; JOINED.
DR EMBL; U83430; AAC53585.1; JOINED.
DR EMBL; U83431; AAC53585.1; JOINED.
DR EMBL; U83432; AAC53585.1; JOINED.
DR EMBL; U83433; AAC53585.1; JOINED.
DR EMBL; U83434; AAC53585.1; JOINED.
DR EMBL; Y16524; CAA76272.1; -.
DR EMBL; Y16525; CAA76272.1; JOINED.
DR EMBL; Y16526; CAA76272.1; JOINED.
DR EMBL; Y16527; CAA76272.1; JOINED.
DR EMBL; Y16528; CAA76272.1; JOINED.
DR EMBL; Y16529; CAA76272.1; JOINED.
DR EMBL; Y16530; CAA76272.1; JOINED.
DR EMBL; Y16531; CAA76272.1; JOINED.
DR EMBL; Y16532; CAA76272.1; JOINED.
DR MGD; MGI:103080; PIGR.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 4.
DR SMART; SM00410; Ig_Like; 2.
DR Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal.
KW SIGNAL 1 18
FT CHAIN 19 771 POTENTIAL.
FT CHAIN 19 611 SECRETORY COMPONENT.
FT DOMAIN 19 645 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 646 668 POTENTIAL.
FT DOMAIN 669 771 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 33 117 IG-LIKE V-TYPE DOMAIN 1.
FT DOMAIN 145 227 IG-LIKE V-TYPE DOMAIN 2.
FT DOMAIN 250 331 IG-LIKE V-TYPE DOMAIN 3.
FT DOMAIN 363 447 IG-LIKE V-TYPE DOMAIN 4.
FT DOMAIN 477 553 IG-LIKE V-TYPE DOMAIN 5.
FT DISULFID 40 110 POTENTIAL.
FT DISULFID 152 220 POTENTIAL.
FT DISULFID 257 324 POTENTIAL.
FT DISULFID 370 440 POTENTIAL.
FT DISULFID 484 546 POTENTIAL.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 159 159 A -> V (IN REF. 1).
FT CONFLICT 396 396 V -> A (IN REF. 1).
FT CONFLICT 620 620 G -> R (IN REF. 1).
FT CONFLICT 771 771 AA: 84998 MW; 78C81302EC710730 CRC64;
SQ SEQUENCE 771 AA: 9.5%; Score 195; DB 1; Length 771;
Query Match 9.5%; Score 195; DB 1; Length 771;
Best Local Similarity 34.5%; Pred. No. 7e-07;
Matches 48; Conservative 21; Mismatches 54; Indels 16; Gaps 4;
QY 10 FLPSGALRILPEVKEGELGSGVTKCPLE-----MHVRIYLCREMAGSGCGPVSTT 65
Db 13 FSGVSTKSPFGQEVSSIEGDSVSITCYPTSVNHRTRKWCROGA-SGMCUTLISSN 71
QY 66 NFIKAEYKGRVTLKQYPRKNLFLVEVTLQTSDSGVACAGAGMTDRGKTOKVTLNVHSE 125
Db 72 GYLSKEYSGRANLINFENNFTVINIEQLTQDDTSYKCGLG-TSRNGLSFDVSLV--- 127
QY 126 YEPSWEQMPETPKWFL 144
Db 128 -----SQVPELPSPDTHV 139
Query Match 9.0%; Score 185; DB 1; Length 769;
Best Local Similarity 34.5%; Pred. No. 3.7e-06;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
RESULT 2
PIGR_RAT
ID PIGR_RAT STANDARD; PRT; 769 AA.
```

```
AC P15083:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Polymorphic-immunoglobulin receptor precursor (Poly-IG receptor) (PIGR)
DE [Contains: Secretory component].
GN PIGR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Banting G., Brake B., Braghetta P., Luzio J.P., Stanley K.K.;
RT "Intracellular targeting signals of polymeric immunoglobulin
RL FEBS Lett. 254:177-183(1989).
CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X15741; CAA33758.1; -.
DR PIR; S05407; S05407.
DR HSSP; P00702; 1JHL.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 4.
DR SMART; SM00410; Ig_Like; 2.
KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 769 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 611 SECRETORY COMPONENT.
FT DOMAIN 19 643 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 644 666 POTENTIAL.
FT DOMAIN 667 769 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 33 117 IG-LIKE V-TYPE DOMAIN 1.
FT DOMAIN 145 227 IG-LIKE V-TYPE DOMAIN 2.
FT DOMAIN 250 331 IG-LIKE V-TYPE DOMAIN 3.
FT DOMAIN 363 447 IG-LIKE V-TYPE DOMAIN 4.
FT DOMAIN 477 553 IG-LIKE V-TYPE DOMAIN 5.
FT DISULFID 40 110 POTENTIAL.
FT DISULFID 152 220 POTENTIAL.
FT DISULFID 257 324 POTENTIAL.
FT DISULFID 370 440 POTENTIAL.
FT DISULFID 484 546 POTENTIAL.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 769 AA: 84798 MW; 5F849303400255A7 CRC64;
Query Match 9.0%; Score 185; DB 1; Length 769;
Best Local Similarity 34.5%; Pred. No. 3.7e-06;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
```


FT	DOMAIN	563	573	EXTRACELLULAR (POTENTIAL).	QY	223	QTPSYNHHTRLHR-----ORALDYGSQSGRE-----GQGFHILPTILGL	262
FT	TRANSMEM	574	592	S4 OF REPEAT II (POTENTIAL).	DB	2049	---SHHHHRRRRDRKQSRLEKGPSLSADMGAPSAVGPGLPPEG-----PT-----	2096
FT	DOMAIN	593	611	S5 OF REPEAT II (POTENTIAL).	QY	263	FLALLGLVVKRAVRRKALSRARRLAVRMRALESSORP-----RGSPPRPSQN	312
FT	TRANSMEM	612	631	EXTRACELLULAR (POTENTIAL).	DB	2097	-----GRRERRRQGRGSRQER-----ROPSSSSSEKORFYSCDRFGGPPKPKPSL	2145
FT	DOMAIN	632	684	S6 OF REPEAT II (POTENTIAL).	QY	313	NIYSACPRAR-----GADAAGTGEAPVPG-----PCAPLPPAP---	349
FT	TRANSMEM	685	709	CYTOPLASMIC (POTENTIAL).	DB	2146	SSHTPTAGTGGPPHPPQSGSVNGSPLLSTSGASTPORGRRRLPOTPLTPRPSITYKT	2205
FT	DOMAIN	710	1151	S1 OF REPEAT III (POTENTIAL).	QY	350	SESPWLHAPSUKTS-----CEYVSLYH-----QPAAMMEDSDSDYI	386
FT	TRANSMEM	1152	1169	EXTRACELLULAR (POTENTIAL).	DB	2206	ANSSPIHFAGAGTSLPAFSPGRSLSEHNALLQRDPLSQPLAPGSRIGSDPYL	2260
FT	DOMAIN	1170	1185	S2 OF REPEAT III (POTENTIAL).	RESULT 12			
FT	TRANSMEM	1186	1205	CYTOPLASMIC (POTENTIAL).	ID	PVR2_MOUSE	STANDARD:	PRT: 530 AA.
FT	DOMAIN	1206	1217	S3 OF REPEAT III (POTENTIAL).	AC	P32507: Q62096;		
FT	TRANSMEM	1218	1236	EXTRACELLULAR (POTENTIAL).	DT	01-OCT-1993 (Rel. 27, Created)		
FT	DOMAIN	1237	1246	S4 OF REPEAT III (POTENTIAL).	DT	16-OCT-2001 (Rel. 40, Last sequence update)		
FT	TRANSMEM	1247	1265	CYTOPLASMIC (POTENTIAL).	DE	01-MAR-2002 (Rel. 41, Last annotation update)		
FT	DOMAIN	1266	1284	S5 OF REPEAT IV (POTENTIAL).	DE	Poliiovirus receptor related protein 2 precursor (Murine herpesvirus entry protein B) (mhvEb) (Nectin 2) (Poliiovirus receptor homolog).		
FT	TRANSMEM	1285	1304	EXTRACELLULAR (POTENTIAL).	GN	PVR2 OR PVS OR PVR OR MPH.		
FT	DOMAIN	1305	1391	S6 OF REPEAT III (POTENTIAL).	OS	Mus musculus (Mouse).		
FT	TRANSMEM	1392	1416	CYTOPLASMIC (POTENTIAL).	OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
FT	DOMAIN	1417	1471	S1 OF REPEAT IV (POTENTIAL).	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
FT	TRANSMEM	1472	1490	EXTRACELLULAR (POTENTIAL).	OX	NCBI_TaxID=10090;		
FT	DOMAIN	1491	1505	S2 OF REPEAT IV (POTENTIAL).	RN	[1]		
FT	TRANSMEM	1506	1525	CYTOPLASMIC (POTENTIAL).	RP	SEQUENCE FROM N.A. (ISOFORM ALPHA).		
FT	DOMAIN	1526	1533	S3 OF REPEAT IV (POTENTIAL).	RX	MEDLINE=92219365; PubMed=1560525;		
FT	TRANSMEM	1534	1552	EXTRACELLULAR (POTENTIAL).	RA	Morrison M.E.; Racaniello V.R.;		
FT	DOMAIN	1553	1563	S4 OF REPEAT IV (POTENTIAL).	RT	"Molecular cloning and expression of a murine homolog of the human poliovirus receptor gene.";		
FT	TRANSMEM	1564	1582	CYTOPLASMIC (POTENTIAL).	RL	J. Virol. 66:2807-2813(1992).		
FT	DOMAIN	1583	1601	S5 OF REPEAT IV (POTENTIAL).	RN	[2]		
FT	TRANSMEM	1602	1621	EXTRACELLULAR (POTENTIAL).	RP	SEQUENCE FROM N.A. (ISOFORM BETA).		
FT	DOMAIN	1622	1683	S6 OF REPEAT IV (POTENTIAL).	RC	STRAIN=C57BL/6; TISSUE=Brain;		
FT	TRANSMEM	1684	1708	CYTOPLASMIC (POTENTIAL).	RX	MEDLINE=94179228; PubMed=8132569;		
FT	DOMAIN	1709	2339	POLY-HIS.	RA	Aoki J., Koike S., Ise I., Sato-Yoshida Y., Nomoto A.;		
FT	TRANSMEM	2050	2054	BINDING TO THE BETA SUBUNIT (BY SIMILARITY).	RT	"Amino acid residues on human poliovirus receptor involved in interaction with poliovirus.";		
FT	DOMAIN	2118	2122	ATP (POTENTIAL).	RL	J. Biol. Chem. 269:8431-8438(1994).		
FT	TRANSMEM	379	396	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).	RN	[3]		
FT	NP_BIND	451	458	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).	RP	CHARACTERIZATION.		
FT	SITE	314	314	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).	RX	MEDLINE=99214397; PubMed=10196354;		
FT	SITE	663	663	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).	RA	Shukla D., Rowe C.L., Dong Y., Racaniello V.R., Spear P.G.;		
FT	SITE	1365	1365	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).	RT	"The murine homolog (Mph) of human herpesvirus entry protein B (HvEb) mediates entry of pseudorabies virus but not herpes simplex virus types 1 and 2.";		
FT	MOD_RES	1719	1719	CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY).	RT	J. Virol. 73:4493-4497(1999).		
FT	CA_BIND	1737	1748	PHOSPHORYLATION (BY CAPK) (POTENTIAL).	CC	!- FUNCTION: RECEPTOR FOR ALPHAPERHESVIRUS (SUCH AS MURINE HSV) ENTRY INTO CELLS.		
FT	CARBOHYD	256	256	BY SIMILARITY.	CC	!- SUBCELLULAR LOCATION: Type I membrane protein.		
FT	CARBOHYD	1563	1563	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA AND BETA (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.		
FT	VARSPLIC	2164	2339	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	!- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, SPLEEN, KIDNEY, HEART AND LIVER.		
FT	MOD_RES	1719	1719	SGSVNGSPLLSTSGASTPGRGRRLPQPLTPRPSITYK	CC	!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.		
FT				TANSSPIHFAAQOTSLPAFSPGRSLSEHNALLQRDPLS	CC	!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.		
FT				QPLAPGRIGSDPYLQRLDSEASVHALPDTLTFEEAVAT	CC	!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.		
FT				NSGRSSRTSYSSLTQSQSHPLRRYPNGVHPTCLGLSSGGRAR	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
FT				HSYHHPDQDHK -> AGSANGFPNTTPCCCTETSPASWPPL	CC			
FT				ALEALALTWSGVTVVRPLSTPLCLRTSLSLRRLWPPTRAAP	CC			
FT				PGLPTRCPP (IN ISOFORM ALPHA-IB-2).	CC			
FT	SEQUENCE	2339 AA;	262494 MW;	17A45C6D1E76B39D CRC64;	CC			
FT	Query Match	5.3%;	Score 109.5;	DB 1; Length 2339;	CC			
FT	Best Local Similarity	20.0%;	Pred. No. 4.6;		CC			
FT	Matches	83;	Conservative	46; Mismatches 135; Indels 151; Gaps 18;	CC			
QY	75	RVTLKQYPRKNLFLVEVTOLETSDGVYACGAGMNTDRGKTOKVTLNVHSEYEPWSWESQ	134		CC			
DB	1894	KATLEQTQPAVLRGARVFLRQKSTSLSSNGGATQNOESGIKESV-----SWGTR	1943		CC			
QY	135	MPETPKWFHLP-----YLFQMPAYASS-----KEVTR-----VTPP	166		CC			
DB	1944	TQDAPAEARPLERGHSTELPVGRSGALAVDVMQSGITIRRGDPGPPQGLCSQSGRAASMP	2003		CC			
QY	167	AQRGKVPVPHSSP---TQIITHPRVSRASSVAGDKPRTFLPSTTASTKISALEGLLKP	222		CC			
DB	2004	PLAFTQPVNDASPMKPSISTLAQREGTICSTTDRP-----PPSOAS-----	2048		CC			


```
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 3.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00784; MYTH4; 2.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 2.
DR SMART: SM00242; MYSC; 1.
DR SMART: SM00139; MYTH4; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00660; BAND_41_1; FALSE_NEG.
DR PROSITE: PS00661; BAND_41_2; FALSE_NEG.
DR PROSITE: PS50057; BAND_41_3; 1.
DR PROSITE: PS50096; IQ; 3; FALSE_NEG.
DR PROSITE: PS50002; SH3; FALSE_NEG.
KW Myosin: ATP-binding; Actin-binding; Coiled coil; Repeat; SH3 domain;
KW Calmodulin-binding; Disease mutation; Deafness.
FT DOMAIN 1 1887 HEAD OR MOTOR DOMAIN.
FT DOMAIN 1888 2029 NECK OR REGULATORY DOMAIN.
FT DOMAIN 2030 3530 TAIL.
FT DOMAIN 1323 1350 COILED COIL (POTENTIAL).
FT DOMAIN 1792 1799 ACTIN-BINDING (POTENTIAL).
FT DOMAIN 1902 1924 IQ 1.
FT DOMAIN 1925 1954 IQ 2.
FT DOMAIN 1955 1976 IQ 3.
FT DOMAIN 2867 2953 SH3.
FT DOMAIN 3206 3443 BAND 4.1-LIKE.
FT NP_BIND 1315 1322 ATP (POTENTIAL).
FT VARIANT 2111 2111 N -> Y (IN DFN83; FAMILY FROM BENGKALA).
FT VARIANT 2113 2113 /FTID-VAR_010303.
FT VARIANT 2113 2113 I -> F (IN DFN83; INDIAN FAMILY).
FT SEQUENCE 3530 AA; 395171 MW; 3D103923D4BCBE4A CRC64;
Query Match 4.9%; Score 101.5; DB 1; Length 3530;
Best Local Similarity 21.3%; Pred. No. 29;
Matches 86; Conservative 47; Mismatches 138; Indels 133; Gaps 21;
QY 14 SGALRIILPEYKVE-----GELGSGVTTCPLPEMHVRIYLCREMAG----- 54
DB 2200 SGAARTLPPTQLEWTATYFKASMALDVGCFNGD-QFSCPVHWSWG-----EEVAGDILRH 2254
QY 55 -----SGTCGVVSTNPIK-AEYKGR-----VTLKQYPRKNLFLVEVNTQLTSDS 99
DB 2255 RGLADGWRGWTVMKNGVQMAELAGHDYVLDLVDLSDLELDFPRQKSYFIVGTE----- 2308
QY 100 GYVACGAGMNTDRGKTQKVTLVNHSEYEPSWEE-----QPMPEPKWFHPLPYLFQMPAY 153
DB 2309 -----GPAASRGPRKPVFGN-----SWDSDEDMSTRPQPE-----HMPKVLSDGY 2350
QY 154 ASSSKFVTRVTTPAQRGKVPVHHSSPTQITHRPVRVSASSVAGDKPRFLDSTASKI 213
DB 2351 SSHNQDNGTEAQRGTA-----TH-----QESDSLGE-----PAVPHKGL 2387
QY 214 SA-LEGELKLPOTPSYNHHTLHRORALDY-----GSGSGREGQGFH-----ILIPT 258
DB 2388 DCYLDLSLFDLP-VLSYG-DADLEKPTAIYRMKGGGQPGGSGSGTETPRRPEPKPIFG 2445
QY 259 ILGLFLALLGLLVKRAVERKALSRA-----RRLAVRMRALESSQRPRGSPRPSQNN 313
DB 2446 LDASTLALQQAFTHKQAVLLARGMTLQATALQQOPLSAALRSIPAEPKPAPEAQPTS--- 2502
QY 314 IYSACPRRACDAAGTG--EAPVPGCAPLPPAPLOVSESPML 355
DB 2503 -----VGTGPPAKPVLVLRATPKPLAPAPLAKAPRL 2532
RESULT 14
ID CCAB_RABIT STANDARD; PRT; 2339 AA.
AC Q05152;
DT 15-JUL-1999 (Rel. 38, Created)
```

15-JUL-1999 (Rel. 38, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
Voltage-dependent N-type calcium channel alpha-1B subunit (Calcium channel, L type, alpha-1 polypeptide isoform 5) (Brain calcium channel III) (BIII).
GN CACNA1B OR CACNL1A5 OR CACH5.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid-9986;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93236885; PubMed=8386525;
RA Fujita Y., Mylnieff M., Dirksen R.T., Kim M.-S., Nildome T., Nakai J., Friedrich T., Iwabe N., Miyata T., Furuichi T., Furutama D., Mikoshiba K., Mori Y., Beam K.G.;
RA "Primary structure and functional expression of the omega-conotoxin-sensitive N-type calcium channel from rabbit brain.";
RL Neuron 10:585-598(1993).
CC -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION, CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1B GIVES RISE TO N-TYPE CALCIUM CURRENTS. N-TYPE CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP AND ARE BLOCKED BY OMEGA-CONOTOXIN-GVIA (OMEGA-CTX-GVIA) AND BY OMEGA-AGATOXIN-IIIA (OMEGA-AGA-IIIA). THEY ARE HOWEVER INSENSITIVE TO DIHYDROPYRIDINES (DHP), AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA). CALCIUM CHANNELS CONTAINING ALPHA-1B SUBUNIT MAY PLAY A ROLE IN DIRECTED MIGRATION OF IMMATURE NEURONS.
CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: WIDESPREAD EXPRESSION THROUGHOUT THE BRAIN. HIGHEST LEVELS IN CORPUS STRIATUM AND MIDBRAIN.
CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC -!- PTM: PHOSPHORYLATED IN VITRO BY CAM-KINASE II, CAPK, PKC AND CGPK (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS FAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D14157; BAA03202.1; -
DR InterPro: IPR002077; Ca_channel.
DR InterPro: IPR002111; Cat_channel_TrlP.
DR InterPro: IPR000636; Cation_chan_non_lig.
DR InterPro: IPR001882; Channel_pore_Ca_Na.
DR Pfam: PF00520; Ion_trans; 4.
DR PRINTS: PR00167; CACHANNEL.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel; Glycoprotein; Repeat; Multigene family;
KW Calcium-binding; Phosphorylation; ATP-binding.
FT REPEAT 82 359 I.
FT REPEAT 469 713 II.
FT REPEAT 1142 1424 III.

Page 13

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR	ENBL; X52491; CAA36731.1; -.	
DR	PIR; S10017; S10017.	
DR	PDB; 4GAT; 28-JAN-98.	
DR	PDB; 5GAT; 28-JAN-98.	
DR	PDB; 6GAT; 28-JAN-98.	
DR	PDB; 7GAT; 28-JAN-98.	
DR	TRANSFAC; T02533; -.	
DR	InterPro; IPR000679; ZnF_GATA.	
DR	Pfam; PF003320; GATA: 1.	
DR	PRINTS; PR00619; GATAZNFINGER.	
DR	SMART; SM00401; znF_GATA: 1.	
DR	PROSITE; PS00344; GATA-ZN_FINGER_1; 1.	
DR	PROSITE; PS00114; GATA-ZN_FINGER_2; 1.	
KW	transcription regulation; Activator; DNA-binding; Zinc-finger;	
KW	Nuclear protein; Nitrate assimilation; 3D-structure.	
FT	ZN_FING 673 697	
FT	GATA-TYPE.	
FT	H-T-H MOTIF (PROBABLE).	
FT	SEQUENCE 876 AA; 94195 MW; 9ADC2273EE536F98 CRC64;	
SO	SEQUENCE	

Query Match	4.9%	Score 100.5;	DB 1;	Length 876;
Best Local Similarity	22.0%	Pred. NO. 6.4;		
Matches	77:	Conservative	47:	Mismatches 137;
				Indels 89;
				Gaps 17;

Qy	96	ESDSG--VYACGAGMNTDRGKTKQKVLNVHSEYEPSEWESQMPETPKWPHLPLYLQ-MPA 152
Db	515	EDDDGNQFSAGGLAMPAEFGDDSIDMNSMAWETSY-----PNSFQSLPA 560
Qy	153	YASS-SKFEV-----RVTTPA---QRGKVPVHHSSPT-TQITHRPVSRASVAGDKPR 202
Db	561	FAAQRKHVITGSADMMDTSENNQGSGLRTHESAASYSEVRNDDQPRQKIA---R 616
Qy	203	TFLPSTTASIKISALEGLLKQFTSYNHHT-----RLHQRALDYGSSQG 246
Db	617	TSSTPTNAQ-----LLROSMQNSHSTSPNTPPEGSLNSAAPSAPGGTKNGEQNG 669
Qy	247	-----REGQGFHILIPITILGLFLLALGLVVKRAVERRKALSRARR 288
Db	670	PTTCTNCFTQTTPLWRNPESQP-----LCNACGLF-LKLHGVV--RPLSLKTDVVIKRNK 722
Qy	289	LAVRMRALESSORPGSPRPQNNIYSACPRARGADAAGTCEAPVPGGA----- 340
Db	723	NSANSLAVGSRVSKKARKNSVQQVTPTAPTSSR-AQSNWTTSEPPAMPGSGRGSGVV 781
Qy	341	PLPPAPLQVSESPWLHAPSLKTSCEYVSLYHQPAAMMEIDSDSDDYINVPA 390
Db	782	PIAAAPPK-SSSAATTPGTNGCGAVOVAPKRRKLEKASDVMDAESPS830

Search completed: October 28, 2002, 17:31:23
Job time : 19.0081 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:40 ; Search time 24.7419 Seconds
(without alignments)
1059.473 Million cell updates/sec

Title: US-09-135-238b-2_COPY_18_253
Perfect score: 1254
Sequence: 1 RLPLPVKEGELGGSVTKC.....HRQRALDYGSQSGREGQGFH 236

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802: *
1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	ID	Description
1	1254	100.0	390	20	AA1980
2	1254	100.0	390	20	AA1981
3	1254	100.0	390	20	AA1982
4	659	52.6	422	22	AAE05349
5	459	36.6	107	20	AA1980
6	319	25.4	97	21	AA1981
7	197	15.7	771	17	AA1982
8	194	15.5	771	22	AA1983
9	185	14.8	769	17	AA1984
10	185	14.8	769	22	AA1985
11	180.5	14.4	607	20	AA1986

12	180.5	14.4	607	20	AAW95601	Human secretory Im
13	180.5	14.4	746	17	AAW03178	Human poly-immunog
14	180.5	14.4	764	22	AA1986	Human polymeric im
15	180.5	14.4	764	22	AA1986	Human polymeric im
16	179.5	14.3	757	17	AAW03179	Bovine poly-immuno
17	179	14.3	757	22	AA1986	Bovine polymeric i
18	176	14.0	733	22	AA1986	Possum polymeric i
19	175.5	14.0	532	19	AAW50033	Human immunity rel
20	166.5	13.3	584	12	AA1986	Truncated poly Ig-
21	166.5	13.3	773	17	AAW03177	Rabbit poly-immuno
22	166.5	13.3	773	22	AA1986	Rabbit poly-immuno
23	156	12.4	584	12	AA1986	Truncated poly Ig-
24	126	10.0	211	21	AA1986	Human immune syste
25	121	9.6	313	22	AA1986	Human DNAX surface
26	119.5	9.5	203	22	AA1986	Novel bone marrow
27	119.5	9.5	205	20	AA1986	Human pIGR-2 prote
28	117.5	9.4	334	19	AAW38336	Human kidney injur
29	111.5	8.9	307	19	AAW38334	Rat kidney injury
30	108.5	8.7	268	22	AAW40441	Human polyptide
31	106	8.5	298	20	AAW40423	Human CMRF-35-H9 r
32	106	8.5	301	20	AAW40422	Human CMRF-35-H9 r
33	105	8.4	164	20	AAW40422	Extracellular doma
34	103.5	8.3	332	20	AAW41717	Human PRO846 prote
35	103.5	8.3	332	21	AAW44273	Human PRO846 (UNQ4
36	103.5	8.3	332	21	AAW27653	Human protein PRO8
37	103.5	8.3	332	21	AAW19580	Human PRO846 used
38	103.5	8.3	332	22	AAW12364	Human PRO846 polyp
39	103.5	8.3	332	22	AAW55300	Human PRO846 prote
40	100.5	8.0	504	22	AAW94522	Human protein sequ
41	100.5	8.0	504	22	AAW50245	Human transcriptio
42	98.5	7.9	230	22	AAW38655	Human polyptide
43	98.5	7.9	231	19	AAW75138	Human secreted pro
44	98.5	7.9	260	21	AAW87353	Human signal pepti
45	98	7.8	493	22	AAW93283	Human polypeptide,

ALIGNMENTS

RESULT 1
AA1986
ID AA1986 standard; Protein: 390 AA.
XX
AC AA1986
XX
DT 17-DEC-1999 (first entry)
XX
DE Human Toso protein sequence.
XX
KW Human: Toso protein; target; drug screening; diagnosis; apoptosis;
KW apoptosis related disease.
XX
OS Homo sapiens.
XX
PN WO9950671-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06945.
XX
PR 30-MAR-1998; 98US-0050861.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Payan D;
XX
DR WPI; 1999-591379/50.
DR N-PSDB; AA25422.
XX
PT Screening agents useful for modulating apoptosis and controlling
PT apoptosis related diseases -
XX
PS Claim 1; Fig 2a; 75pp; English.

XX The present invention describes a method of Screening for a bioactive
 CC agent capable of binding a Toso protein. Also described a methods for:
 CC (1) screening a bioactive agent capable of modulating activity of a
 CC Toso cell-surface receptor, comprising adding a candidate bioactive
 CC agent to a cell comprising a recombinant Toso nucleic acid, and
 CC exposing the cells to an apoptotic agent that will induce apoptosis;
 CC (2) modulating apoptosis comprising administering an exogenous
 CC compound that binds Toso, to a cell; (3) identifying a cell containing
 CC a mutant Toso gene, comprising determining it's sequence; (4)
 CC identifying the Toso genotype, comprising determining the sequence of
 CC at least one Toso gene; and (5) diagnosing an apoptosis related
 CC condition, comprising measuring activity of Toso in a tissue, and
 CC comparing to the activity from non-affected individual's tissue, where
 CC a reduced activity of the patient indicates risk of an apoptosis related
 CC condition. The methods are useful for identifying agents capable of
 CC diagnosing and treating apoptosis related disease, their use for
 CC modulating apoptosis, and methods for diagnosing the disease state.
 CC The present sequence represents the human Toso protein for use in
 CC methods from the invention.

XX Sequence 390 AA;

Query Match 100.0%; Score 1254; DB 20; Length 390;
 Best Local Similarity 100.0%; Pred. No. 9.3e-111;
 Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGTWSTTFNFIKAEYKGRVT 60
 |||||||
 Db 18 RILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGTWSTTFNFIKAEYKGRVT 77
 |||||||

QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSWEQPMPE 120
 |||||||

Db 78 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSWEQPMPE 137
 |||||||

QY 121 TPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQTTHRPVRSASSVA 180
 |||||||

Db 138 TPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQTTHRPVRSASSVA 197
 |||||||

QY 181 GDKPRTFLPSTTASIKSALGELLPQTPSYNHHTRLHRQALDYGSGSGREGQGFH 236
 |||||||

Db 198 GDKPRTFLPSTTASIKSALGELLPQTPSYNHHTRLHRQALDYGSGSGREGQGFH 253
 |||||||

RESULT 2

AAI17496

ID AAI17496 standard; Protein; 390 AA.

XX AAI17496;

XX 03-AUG-1999 (first entry)

XX Human Toso protein.

XX Toso protein; tumour necrosis factor mediated apoptosis inhibition;
 KW TNF mediated apoptosis; T cell overactivity; autoimmune disease;
 KW Sjogrens connective tissue disorder; transplant rejection; cancer.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..17

FT /label= signal

FT Protein 18..390

FT /label= Toso

FT Region 254..272

FT /label= transmembrane_region

XX W09525832-A1.

XX 27-MAY-1999.

XX 16-NOV-1998; 98WO-US24391.

XX 17-AUG-1998; 98US-0135238.
 PR 17-NOV-1997; 97US-0066063.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Hitoshi Y, Nolan GP;
 XX
 DR WPI; 1999-338007/28.
 DR N-PSDB; AAX76123.
 XX
 PT DNA encoding Toso, a protein having inhibitory effects on TNF
 mediated apoptosis
 XX
 PS Claim 20; Fig 2a; 70pp; English.

XX The present sequence is a Toso protein (I). (I) has anti-apoptotic
 CC and cytostatic activity. Toso (named after a Japanese liquor that is
 CC drunk on New Year's Day to celebrate long life and eternal youth) most
 CC likely acts by induction of cFLIP expression which inhibits caspase-8
 CC processing. Recombinant (I) can be used to modulate apoptosis in a cell
 CC or to treat an apoptosis related condition in a mammal. Apoptosis
 CC related conditions can also be treated by administration of the Toso
 CC protein or antibody. Apoptosis related or mediated conditions that can
 CC be treated include diseases characterized by T cell overactivity, e.g.
 CC Sjogrens connective tissue disorder, autoimmune diseases, diseases where
 CC T cells actively destroy cells, including transplant rejection and
 CC conditions where cells of any kind that are not dying express Toso
 CC appropriately, e.g. cancer of T or B cell origin (where increased
 CC apoptosis would be appropriate).

XX Sequence 390 AA;

Query Match 100.0%; Score 1254; DB 20; Length 390;
 Best Local Similarity 100.0%; Pred. No. 9.3e-111;
 Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGTWSTTFNFIKAEYKGRVT 60
 |||||||

Db 18 RILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGTWSTTFNFIKAEYKGRVT 77
 |||||||

QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSWEQPMPE 120
 |||||||

Db 78 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSWEQPMPE 137
 |||||||

QY 121 TPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQTTHRPVRSASSVA 180
 |||||||

Db 138 TPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQTTHRPVRSASSVA 197
 |||||||

QY 181 GDKPRTFLPSTTASIKSALGELLPQTPSYNHHTRLHRQALDYGSGSGREGQGFH 236
 |||||||

Db 198 GDKPRTFLPSTTASIKSALGELLPQTPSYNHHTRLHRQALDYGSGSGREGQGFH 253
 |||||||

RESULT 3

AAI05001

ID AAI05001 standard; Protein; 390 AA.

XX AAI05001;

XX 16-JUN-1999 (first entry)

XX Human PIGRL-1 protein sequence.

XX PIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;
 KW X-linked Severe Combined Immunodeficiency; XSCID; IgA deficiency;
 KW diagnosis; therapy.

XX Homo sapiens.

XX EP905238-A2.

XX 31-MAR-1999.

XX PF 14-AUG-1998; 98EP-0306487.
XX PR 30-OCT-1997; 97US-0961564.
XX PR 25-AUG-1997; 97US-0056935.
XX PA (SMIK) SMITHKLINE BEECHAM CORP.
XX PI Sweet RW, Truneh A, Wu S;
XX PI WPI; 1999-192666/17.
XX DR N-PSDB; AAX28178.
XX PT New polypeptides encoding PIGRL-1 useful for treating diseases such
XX as X-linked Severe Combined Immunodeficiency
XX Claim 11; Page 7; 26pp; English.
XX This sequence is the human PIGRL-1 protein of the invention.
CC Autoimmune diseases involving altered expression or activity of PIGRL-1
CC may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined
CC Immunodeficiency (XSCID) and IgA deficiency. These diseases can be
CC diagnosed or susceptibility to them predicted by: (1) determining whether
CC there is a mutation in the genomic copy of the gene encoding PIGRL-1; or
CC (2) measuring the amount of PIGRL-1 in a sample derived from the patient.
CC Patients deficient in PIGRL-1 can be treated by administering either the
CC PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient.
CC Patients with excessive expression or activity of PIGRL-1 can be treated
CC by administering an antagonist of PIGRL-1, an antisense nucleic acid
CC molecule which inhibits the expression of PIGRL-1 or administering
CC sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can
CC be used to identify its agonists by contacting a cell expressing PIGRL-1
CC with a candidate compound in the presence of a signal system and noting
CC the candidate as an agonist if a signal is produced. The same method can
CC be used to identify antagonists of PIGRL-1 but the presence of an
CC antagonist is indicated by a decrease in production of the signal.
CC Antibodies against PIGRL-1 may be used to isolate or identify clones
CC expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to
CC identify chromosomal mutations in the gene encoding PIGRL-1 in patients.
CC This information may then be correlated with the incidence of autoimmune
CC disease in those patients to identify whether the mutation causes the
XX disease.
XX SQ Sequence 390 AA;
Query Match 100.0%; Score 1254; DB 20; Length 390;
Best Local Similarity 100.0%; Pred. No. 9.3e-111;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RILPEVKVEGELGGSVTIKCPLEPMHVRIYLCREMAGSGTCGTVVSTTNFIKAEYKGRVT 60
DB 18 RILPEVKVEGELGGSVTIKCPLEPMHVRIYLCREMAGSGTCGTVVSTTNFIKAEYKGRVT 77
QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSWEQPMPE 120
DB 78 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSWEQPMPE 137
QY 121 TPKNFHLPLFQMPAYASSKSFVTRVTPAQRGKVPVPHHSSPTQTTHRPRVSRASSVA 180
DB 138 TPKNFHLPLFQMPAYASSKSFVTRVTPAQRGKVPVPHHSSPTQTTHRPRVSRASSVA 197
QY 181 GDKPRFTLPSTASKISALEGLLPQTPSYNHHTRLHRQALDYGSGSGREGQGQFH 236
DB 198 GDKPRFTLPSTASKISALEGLLPQTPSYNHHTRLHRQALDYGSGSGREGQGQFH 253
RESULT 4
ID AAE05349
XX AAE05349 standard; Protein; 422 AA.
XX AC AAE05349;
XX DT 12-SEP-2001 (first entry)

XX DE Mouse Toso protein.
XX KW Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;
KW wound healing; immune response; vaccine; cancer; asthma; allergy;
KW cell trafficking; therapy; secreted protein; Fas-induced apoptosis;
KW Toso.
XX OS Mus sp.
XX PN WO200148192-A1.
XX PD 05-JUL-2001.
XX PF 21-DEC-2000; 2000WO-NZ00256.
XX PR 23-DEC-1999; 99US-0171678.
XX PR 28-NOV-2000; 2000US-0724864.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Watson JD, Murison JG;
XX WPI; 2001-425665/45.
XX DR N-PSDB; AAD10117.
XX Novel isolated polypeptide useful to isolate corresponding interacting
XX proteins or other compounds, to quantitatively determine levels of
XX interacting proteins or other compounds, and as therapeutic target -
XX Claim 6; Page 78-79; 101pp; English.
XX The patent discloses novel polynucleotides and their corresponding
XX proteins which play a major role in induction of growth, cell migration
XX and proliferation, cell-cell interaction and the differentiation of
XX tissue-specific cells. These proteins are important in the maintenance
XX of tissue integrity and thus are important in wound healing. They are
XX useful in various assays to determine the biological activity, to raise
XX antibodies, to isolate corresponding interacting proteins or other
XX compounds, to quantitatively determine levels of interacting proteins or
XX other compounds, and as therapeutic target in a whole range of disease
XX states. Compositions comprising the novel proteins of the invention are
XX useful for treating mammalian disorders. Polynucleotides of the invention
XX are useful in genome and physical mapping, in positional cloning of
XX genes, to tag or identify an organism or its reproductive material (as
XX non-disruptive tags for marking organisms), and for the diagnosis and
XX treatment of mammalian diseases which is the consequence of inappropriate
XX expression of kinase genes. They are useful for promoting immune response
XX as part of a vaccine or anti-cancer treatment, as target for cancer
XX treatment, as immunoregulatory and anti-inflammatory molecule, as
XX diagnostic for specific types of cancer and for development of an
XX anti-cancer treatment, and as a target for antagonists in the treatment
XX of diseases such as asthma and allergy. They are also useful to inhibit
XX or enhance the activity of the soluble molecule that binds proteins of
XX the invention, for tissue and neural regeneration, to promote or block
XX cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.
XX The present sequence is Toso, a secreted protein from mouse. Toso is
XX a cell surface, specific regulator of Fas-induced apoptosis in T-cells.
XX SQ Sequence 422 AA;
Query Match 52.6%; Score 659; DB 22; Length 422;
Best Local Similarity 53.7%; Pred. No. 3.5e-54;
Matches 130; Conservative 34; Mismatches 70; Indels 8; Gaps 4;
QY 1 RILPEVKVEGELGGSVTIKCPLEPMHVRIYLCREMAGSGTCGTVVSTTNFIKAEYKGRVT 60
DB 18 RVLPEVQLNVWEGSIIIECPQLQHVRYLCRQAKPGICSTVVSNT-FVKKEYERRY 76
QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLNVHSEY-EPSEWQPM 119
DB 77 LTPCLDKLFLVEMTQLTENDGGIYACGVGMKDKGKTQKITLNVHNEYPEPFWEDEWTS 136

QY	120	ETPKWFLPYLFQMP-----AYASSKFKVTRVTTPAQKGKVPVPVHHSSPTTQTITHRPRVS	174
Dd	137	ERPRWLHRLFQLHQMWLHGSHPSGSGVIKTATVTTSPASKTEAPPVHPQSPSITSVTQHPRVY	196
QY	175	RASSVAGDRPRTFLSTTASIKSLALEGLLKQPOTPSYNHHTLRHORALDYGSGREGOG	234
Dd	197	RAFSVSATKSPALLATTASTSQQA-IRPLEASYSHTRLHEQTRHGHGYREDRG	255
QY	235	FH 236	
Dd	256	LH 257	
RESULT 5			
AAI05002		AAI05002 standard; peptide; 107 AA.	
XX	AC	AAI05002;	
DT	16-JUN-1999	(first entry)	
DE	Human PIGRL-1 protein sequence fragment.		
KW	PIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;		
KW	X-linked Severe Combined Immunodeficiency; XSCID; IgA deficiency;		
KW	diagnosis; therapy.		
OS	Homo sapiens.		
PN	EP905238-A2.		
XX	31-MAR-1999.		
PD			
Pf	14-AUG-1998;	98BP-0306487.	
PR	30-OCT-1997;	97US-0961564.	
PR	25-AUG-1997;	97US-0056935.	
PA	(SMIK) SMITHKLINE BEECHAM CORP.		
XX	Sweet RW, Truneh A, Wu S;		
XX	WPI; 1999-192666/17.		
DR	N-PSDB; AAX28179.		
XX	New polypeptides encoding PIGRL-1 useful for treating diseases such		
PT	as X-linked Severe Combined Immunodeficiency		
PS	Disclosure; Page 9; 26pp; English.		
CC	This sequence is a fragment of the human PIGRL-1 of the invention.		
CC	Autoimmune diseases involving altered expression or activity of PIGRL-1		
CC	may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined		
CC	Immunodeficiency (XSCID) and IgA deficiency. These diseases can be		
CC	diagnosed or susceptibility to them predicted by: (1) determining whether		
CC	there is a mutation in the genomic copy of the gene encoding PIGRL-1; or		
CC	(2) measuring the amount of PIGRL-1 in a sample derived from the patient.		
CC	Patients deficient in PIGRL-1 can be treated by administering either the		
CC	PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient.		
CC	Patients with excessive expression or activity of PIGRL-1 can be treated		
CC	by administering an antagonist of PIGRL-1, an antisense nucleic acid		
CC	molecule which inhibits the expression of PIGRL-1 or administering		
CC	sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can		
CC	be used to identify its agonists by contacting a cell expressing PIGRL-1		
CC	with a candidate compound in the presence of a signal system and noting		
CC	the candidate as an agonist if a signal is produced. The same method can		
CC	be used to identify antagonists of PIGRL-1 but the presence of an		
CC	antagonist is indicated by a decrease in production of the signal.		
CC	Antibodies against PIGRL-1 may be used to isolate or identify clones		
CC	expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to		
CC	identify chromosomal mutations in the gene encoding PIGRL-1 in patients.		
CC	This information may then be correlated with the incidence of autoimmune		
CC	disease in those patients to identify whether the mutation causes the		

CC disease.
XX
SQ Sequence 107 AA;

Query Match 36.6%; Score 459; DB 20; Length 107;
Best Local Similarity 100.0%; Pred.No. 5,7e-36;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILPEVKVEGLGGSVTKCPPEMHVRILCREMAGSGTCGTIVYSTNFIKAIEYGRVT 60
|||||
Db 18 RILPEVKVEGLGGSVTKCPPEMHVRILCREMAGSGTCGTIVYSTNFIKAIEYGRVT 77
|||||
QY 61 LKOYPRKNLFIVEVTQLTESDSGVACG 88
|||||
Db 78 LKOYPRKNLFIVEVTQLTESDSGVACG 105
|||||

RESULT 6
AAI65401
ID AAI65401 standard; Protein; 97 AA.
XX AC AAI65401;
XX
DT 01-FEB-2000 (first entry)
XX
DE Human 5' EST related polypeptide SEQ ID NO:1562.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification.
XX
OS Homo sapiens.
XX
PN WO9953051-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99WO-IB00712.
XX
PR 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-038446/03.
DR N-PSDB; AAZ43015.
XX
PT Novel secreted protein 5' expressed sequence tag sequences used in
diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 3; Page 820-821; 837pp; English.
XX
CC AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
sequences, corresponding to human secreted proteins. AAI64651 to
AAI65438 represent the EST-related proteins corresponding to AAZ42265 to
AAZ43052. The 5' ESTs can be used for producing secreted human gene
products. They can be used to identify and isolate 5' untranslated
regions (UTRs) and upstream regulatory regions which control the
location, development stage, rate, and quantity of protein synthesis, as
well as stability of mRNA. The ESTs are also useful as probes for
chromosome mapping, and to obtain full length cDNA clones. The ESTs can
also be used in forensic procedures to identify individuals, or in
diagnostic procedures to identify individuals having genetic diseases
resulting from abnormal gene expression. The products may also be used in
gene therapy protocols. The nucleic acids encoding signal peptides can be
used for directing extracellular secretion of a polypeptide or the
insertion of a polypeptide into a membrane, or importing a polypeptide
into a cell. The proteins encoded by the EST sequences may be useful in
treating a variety of human conditions. Secreted proteins have
therapeutic value, and the identification of new secreted proteins is

PT New ligands binding to a specific region of a polymeric immunoglobulin
 PT receptor, useful for transporting therapeutic or diagnostic
 PT compositions into or across cells expressing pigR e.g. in drug delivery
 XX
 PS Disclosure; Fig 1; 102pp; English.
 XX
 CC The invention provides ligands that bind specifically to a region of an
 CC animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves
 CC to produce a stalk region remaining attached to the cell and a secretory
 CC component existing in the organ of interest in several forms. The ligands
 CC do not bind to the stalk or the most abundant form of the secretory
 CC component present in the organ under physiological conditions. The
 CC ligands are useful for transporting therapeutic or diagnostic
 CC compositions into or across cells expressing pigR, useful to introduce
 CC or transport ligands such as antibodies and/or to deliver biologically
 CC active components such as proteins, nucleic acids or detectable labels.
 CC They are used to deliver therapeutic compositions to mucosal surfaces.
 CC such as the gastro-intestinal tract, respiratory system etc. in humans.
 CC They are also useful to label cells expressing pigR, e.g. to distinguish
 CC epithelial cells from a mixed cell population in pathology studies or to
 CC aid in carcinoma diagnosis (since pigR expression is reduced in
 CC carcinomas relative to normal epithelium). They can also be used to
 CC deliver veterinary compositions, especially in mammals such as farm,
 CC domestic or wild mammals or birds e.g. birds reared for human
 CC consumption. The present sequence represents a mouse pigR sequence.
 XX
 XX Sequence 771 AA;

Query Match 15.5%; Score 194; DB 22; Length 771;
 Best Local Similarity 36.1%; Pred. No. 1.2e-09;
 Matches 43; Conservative 20; Mismatches 40; Indels 16; Gaps 4;
 QY 13 GGSVTIKCPLE----MHVRIYLCREMAGSGTCGTVTSTNFIRAEYKGRVTLKQYPRKN 68
 Db 33 GDSVITCYTPDTSVNRHTRKYWCROGA-SGMCTLLISSNGYLSKYSGRANLINFENN 91
 QY 69 LFLVEVTLQTSDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSWEEQPMPTPKWFHL 127
 Db 92 TFVINIHLQTDGTSYKCGLG-TSNRGLSFDVSLEV-----SQVPELPDTHV 139

RESULT 9
 AAW03181
 ID AAW03181 standard; Protein; 769 AA.
 AC AAW03181;
 XX
 DT 24-FEB-1997 (first entry)
 XX
 DE Rat poly-immunoglobulin receptor.
 XX
 KW heavy chain; antigen binding domain; protection; mutants;
 KW mucosal; environment; gastrointestinal; passive; immunisation;
 KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
 KW poly; sorbinus.
 XX

OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT Region 13..45
 FT /note= "putative immunoglobulin binding residues
 FT of domain I"
 FT
 FT Domain 1..120
 FT /label= domain_I
 FT Domain 110..230
 FT /label= domain_II
 FT Domain 210..340
 FT /label= domain_III
 FT Domain 320..450
 FT /label= domain_IV
 FT Domain 440..550

FT /label= domain_v
 FT 550..606
 FT /note= "external portions of domain VI"
 FT 550..627
 FT /note= "external portions of domain VI"
 FT 625..660
 FT /label= transmembrane_segment
 FT 650..769
 FT /label= intracellular_portion
 XX
 PN W09621012-A1.
 XX
 PD 11-JUL-1996.
 XX
 PF 27-DEC-1995; 95WO-US16889.
 XX
 PR 04-MAY-1995; 95US-0434000.
 PR 30-DEC-1994; 94US-0367395.
 XX
 PA (PLAN-) PLANT BIOTECHNOLOGY INC.
 PA (UNWE-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.
 PA (PLAN-) PLANET BIOTECHNOLOGY INC.
 XX
 PI Hiatt AC, Lehner T, Ma JKC;
 XX
 DR WPI; 1996-333987/33.
 DR N-PSDB; AAT31291.
 XX
 XX Immunoglobulin and protection protein complex and its prodn. in
 PT plants - useful for passive immunisation against mucosal antigens,
 PT esp. against S. mutans and S. sorbinus to prevent dental caries
 XX
 PS Disclosure; Pages 123-127; 152pp; English.
 XX
 CC The present sequence is the rat poly-immunoglobulin (Ilg)
 CC receptor, a portion of which corresp. to residues 1-627, pref.
 CC 1-606 or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,
 CC 440-550, 550-606 or 550-627 comprises a protection protein (PP).
 CC The Ig of the invention comprises a PP as above in association with
 CC an Ig derived heavy chain, having at least a portion of an antigen
 CC (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.
 CC gastrointestinal, environments, therefore enhancing its
 CC effectiveness in passively immunising animals against mucosal
 CC pathogens. The Ag binding domain is specifically derived from the
 CC Guy's 13 antibody, and the Ig can be used to prevent dental caries
 CC by binding, e.g. Streptococcus mutans serotypes c, e and f, or
 CC S. sorbinus serotypes d and g.
 XX
 SQ Sequence 769 AA;

Query Match 14.8%; Score 185; DB 17; Length 769;
 Best Local Similarity 34.5%; Pred. No. 8.5e-09;
 Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
 QY 13 GGSVTIKCPLE----MHVRIYLCREMAGSGTCGTVTSTNFIRAEYKGRVTLKQYPRKN 68
 Db 33 GNSVITCYTPDTSVNRHTRKYWCROGA-NGYCATLISSNGYLSKYSGRASLINFENN 91
 QY 69 LFLVEVTLQTSDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSWEEQPMPTPKWFHL 127
 Db 92 TFVINIHLQTDGTSYKCGLG-TTNRGLSFDVSLEV-----SQVPELPDTHV 139

RESULT 10
 AAG65697
 ID AAG65697 standard; protein; 769 AA.
 XX
 AC AAG65697;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE Rat polymeric immunoglobulin receptor (pIgR) sequence.
 XX

KW PolymERIC immunoglobulin receptor; pIgR; ligand; therapeutic;
XX carcinoma diagnosis; veterinary; rat.
OS Rattus sp.

XX WO200172846-A2.

XX PD 04-OCT-2001.

XX PF 26-MAR-2001; 2001WO-US09699.

XX PR 27-MAR-2000; 2000US-192197P.

XX PR 27-MAR-2000; 2000US-192198P.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Mostov KE, Chapin SJ, Richman-Eisenstat J;

XX PS WPI; 2001-611619/70.

XX DR New ligands binding to a specific region of a polymeric immunoglobulin

XX PT receptor, useful for transporting therapeutic or diagnostic

XX PT compositions into or across cells expressing pIgR e.g. in drug delivery

XX PT -

XX PS Disclosure; Fig 1; 102pp; English.

XX CC The invention provides ligands that bind specifically to a region of an

XX CC animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves

XX CC to produce a stalk region remaining attached to the cell and a secretory

XX CC component existing in the organ of interest in several forms. The ligands

XX CC do not bind to the stalk or the most abundant form of the secretory

XX CC component present in the organ under physiological conditions. The

XX CC ligands are useful for transporting therapeutic or diagnostic

XX CC compositions into or across cells expressing pIgR, useful to introduce

XX CC or transport ligands such as antibodies and/or to deliver biologically

XX CC active components such as proteins, nucleic acids or detectable labels.

XX CC They are used to deliver therapeutic compositions to mucosal surfaces

XX CC such as the gastro-intestinal tract, respiratory system etc. in humans.

XX CC They are also useful to label cells expressing pIgR, e.g. to distinguish

XX CC epithelial cells from a mixed cell population in pathology studies or to

XX CC aid in carcinoma diagnosis (since pIgR expression is reduced in

XX CC carcinomas relative to normal epithelium). They can also be used to

XX CC deliver veterinary compositions, especially in mammals such as farm,

XX CC domestic or wild mammals or birds e.g. birds reared for human

XX CC consumption. The present sequence represents a rat pIgR sequence.

XX CC

XX SQ Sequence 769 AA;

Query Match 14.8%; Score 185; DB 22; Length 769;

Best Local Similarity 34.5%; Pred. No. 8.5e-09;

Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 13 GGSVTIKPLPE----MHVRIYLCREMAGSGTCGTGVSTTNFIFAKEYKGRVTLKQYPRKN 68

DB 33 GNSVITCYPTDSVNRHTRKYWCROGA-NGYCATLISSNGLSKEYSGRASLINFPPNS 91

QY 69 LFLVETVLTESDSGVYACGAGMTDRCKTKVTLNVHSEYEPSPWEQPMPTPKWPHL 127

DB 92 TVVINIAHTQEDTGSYKCGLG-TTNRGLFFDVSLEV-----SQVPEFPNDTHV 139

RESULT 11

ID AAY34099 standard; Protein; 607 AA.

XX AAY34099;

XX AC AAY34099;

XX DT 20-DEC-1999 (first entry)

XX DE Partial amino acid sequence of plasmid pSHUSC.

XX KW Multimeric protein; immunoglobulin; receptor-ligand complex;

KW hetero-dimeric receptor; trimeric G protein; transgenic.

XX Synthetic.

XX PN WO9949024-A2.

XX PD 30-SEP-1999.

XX PF 24-MAR-1999; 99WO-US06506.

XX PR 25-MAR-1998; 98US-0079249.

XX PR (PLAN-) PLANET BIOTECHNOLOGY INC.

XX PA Wycoff KL, Jaiswal SK;

XX PI WPI; 1999-580446/49.

XX DR N-PSDB; AA222290.

XX PT Producing heterologous multimeric proteins in plants, transformed with

XX PT several plasmids expressing polypeptide components, particularly for

XX PT immunoglobulins -

XX PT

XX PS Example 1; Fig 8; 42pp; English.

XX CC The invention relates to a method for producing heterologous, multimeric

XX CC proteins in plant cells. The method comprises: (a) transforming the cells

XX CC with several naked plasmids each encoding some, but not all, of the

XX CC polypeptide components of the multimeric proteins, and together providing

XX CC all the polypeptide components; and (b) culturing the cells. The method

XX CC is used to produce biologically active multimeric proteins particularly

XX CC immunoglobulins, receptor-ligand complexes, homo- or hetero-dimeric

XX CC receptors, or trimeric G proteins. This method provides properly

XX CC associated and assembled multimeric proteins in a fast and efficient

XX CC process, without the need to cross plants expressing single component of

XX CC the protein. Transgenic plants containing adjacent and stably integrated

XX CC plasmids, and their progeny can also express the multimeric proteins. The

XX CC present sequence represents the partial amino acid sequence of the

XX CC plasmid pSHUSC.

XX CC

XX SQ Sequence 607 AA;

Query Match 14.4%; Score 180.5; DB 20; Length 607;

Best Local Similarity 38.7%; Pred. No. 1.7e-08;

Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;

QY 4 PEVKVEGELGGSVTIKPLP----EMHVRIYLCREMAGSGTCGTGVSTTNFIFAKEYKGRV 59

DB 25 PE-EVNSVEGNSVITCYPTDSVNRHTRKYWCROGA-CITLISSEGVSSKYAGRA 82

QY 60 TLKQYPRKNLFLVEVTQLTESDSGVYACGAGMTDRCKTKOKVTLNV 105

DB 83 NLINFPENGTFFVNNIAQLSODDSGRYKCGLGINS-RGLSFDVSLEV 127

RESULT 12

ID AAW95601 standard; Protein; 607 AA.

XX AAW95601;

XX AC AAW95601;

XX DT 08-JUN-1999 (first entry)

XX DE Human secretory Immunoglobulin A component.

XX KW Immunoglobulin A; secretory; component; IgA; human; treatment;

XX KW prevention; infection; HIV; AIDS; cold; flu; virus;

XX KW human immunodeficiency virus; respiratory syncytial virus.

XX XX Homo sapiens.

XX OS Homo sapiens.

XX PN WO9857993-A1.

XX XX

RESULT 14

AAG65695

ID AAG65695 standard; protein; 764 AA.

XX

AC AAG65695;

XX

DT 07-JAN-2002 (first entry)

XX

DE Human polymeric immunoglobulin receptor (pIgR) sequence.

XX

KW Polymeric immunoglobulin receptor; pIgR; ligand; therapeutic;

XX

KW carcinoma diagnosis; veterinary; human.

XX

OS Homo sapiens.

XX

FH Key

FT Peptide

FT 487..603

FT /note="peptide to which a ligand binds to (claim 8)";

FT

FT Peptide

FT 487..607

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 487..611

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 487..615

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 487..618

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 520..607

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 520..611

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 520..615

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 520..618

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 544..611

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 544..615

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 544..618

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 560..607

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 560..611

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 560..615

FT /note="peptide to which a ligand binds to"

FT

FT Peptide

FT 560..618

FT /note="peptide to which a ligand binds to"

XX

PR 27-MAR-2000; 2000US-192197P.

PR

XX 27-MAR-2000; 2000US-192198P.

XX

PA (REGC) UNIV CALIFORNIA.

XX

PI Mostov KE, Chapin SJ, Richman-Eisenstat J;

XX

DR WPI; 2001-611619/70.

XX

PT New ligands binding to a specific region of a polymeric immunoglobulin receptor, useful for transporting therapeutic or diagnostic compositions into or across cells expressing pIgR e.g. in drug delivery

PT

PS Disclosure; Fig 1; 102pp; English.

XX

CC

The invention provides ligands that bind specifically to a region of an animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves to produce a stalk region remaining attached to the cell and a secretory component existing in the organ of interest in several forms. The ligands do not bind to the stalk or the most abundant form of the secretory component present in the organ under physiological conditions. The ligands are useful for transporting therapeutic or diagnostic compositions into or across cells expressing pIgR, useful to introduce or transport ligands such as antibodies and/or to deliver biologically active components such as proteins, nucleic acids or detectable labels. They are used to deliver therapeutic compositions to mucosal surfaces. Such as the gastro-intestinal tract, respiratory system etc. in humans. They are also useful to label cells expressing pIgR, e.g. to distinguish epithelial cells from a mixed cell population in pathology studies or to aid in carcinoma diagnosis (since pIgR expression is reduced in carcinomas relative to normal epithelium). They can also be used to deliver veterinary compositions, especially in mammals such as farm, domestic or wild mammals or birds e.g. birds reared for human consumption. The present sequence represents a human pIgR sequence.

SQ

Sequence 764 AA;

Query Match 14.4%; Score 180.5; DB 22; Length 764;
Best Local Similarity 38.7%; Pred. NO. 2.3e-08;
Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;

QY

4 PEVKVEGELGGSVTIKCLPLP---EMHVRIYLCREMAGSGTGVVSTTNFIKAEYKGRV 59

DB

25 PE-EVNSVEGNSVITCYPPPTSVNRHTRKVKWCRGARGG-CITLISSEGVSSKIAGRA 82

QY

60 TLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTLMV 105

DB

83 NLTNFPENGTFVNNIAQLSQDDSGRYKGLGINS-RGLSFDVSLV 127

RESULT 15

AAG65711

ID AAG65711 standard; protein; 764 AA.

XX

AC AAG65711;

XX

DT 07-JAN-2002 (first entry)

XX

DE Human polymeric immunoglobulin receptor (pIgR) sequence.

XX

KW Polymeric immunoglobulin receptor; pIgR; ligand; therapeutic;

XX

KW carcinoma diagnosis; veterinary; human.

XX

OS Homo sapiens.

XX

PN WO200172846-A2.

XX

PD 04-OCT-2001.

XX

PF 26-MAR-2001; 2001WO-US09699.

XX

PR 27-MAR-2000; 2000US-192197P.
PR 27-MAR-2000; 2000US-192198P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
XX Mostov KE, Chapin SJ, Richman-Eisenstat J;
XX
XX WPI; 2001-611619/70.
XX
XX
XX New ligands binding to a specific region of a polymeric immunoglobulin
PT receptor, useful for transporting therapeutic or diagnostic
PT compositions into or across cells expressing pigR e.g. in drug delivery
PT
XX
XX Disclosure; Fig 2; 102pp; English.
XX
XX The invention provides ligands that bind specifically to a region of an
CC animal cell polymeric immunoglobulin receptor (pigR). The pigR cleaves
CC to produce a stalk region remaining attached to the cell and a secretory
CC component existing in the organ of interest in several forms. The ligands
CC do not bind to the stalk or the most abundant form of the secretory
CC component present in the organ under physiological conditions. The
CC ligands are useful for transporting therapeutic or diagnostic
CC compositions into or across cells expressing pigR, useful to introduce
CC or transport ligands such as antibodies and/or to deliver biologically
CC active components such as proteins, nucleic acids or detectable labels.
CC They are used to deliver therapeutic compositions to mucosal surfaces.
CC such as the gastro-intestinal tract, respiratory system etc. in humans.
CC They are also useful to label cells expressing pigR, e.g. to distinguish
CC epithelial cells from a mixed cell population in pathology studies or to
CC aid in carcinoma diagnosis (since pigR expression is reduced in
CC carcinomas relative to normal epithelium). They can also be used to
CC deliver veterinary compositions, especially in mammals such as farm,
CC domestic or wild mammals or birds e.g. birds reared for human
CC consumption. The present sequence represents a human pigR sequence.
XX
XX Sequence 764 AA;
SQ

Query Match 14.4%; Score 180.5; DB 22; Length 764;
Best Local Similarity 38.7%; Pred. No. 2.3e-08;
Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;
Qy 4 PEVKYEGELGSGVTIKCLP-----EMHVRIYLORENAGSGTCGTWSTTFIKAEYKGRV 59
Db 25 PE-EVNSVEGNSVITCYPTPTSVNRHTRKIWCQARGG-CITLISSEGYVSSKYAGRA 82
Qy 60 TLKQYPRKNLFLVEVTQLTSDSGVIYACGAGMNTDRGKTQKVTLVN 105
Db 83 NLTFNFPNGTFVWNIQLSQDDSGRYKGLGINS-RGLSFDVSLEV 127

Search completed: October 28, 2002, 17:30:54
Job time : 27.7419 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 35.121 Seconds
(without alignments)
1921.018 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLWPLYFLPVSGALRIL.....HQRAMMEDSDSDYINVPA 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues 562222
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL19.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2047	99.6	390	060667	Q0667 homo sapien
2	1157	56.3	422	11 Q9D8T1	Q9d8t1 mus musculus
3	186	9.1	534	4 Q96SA2	Q96sa2 homo sapien
4	184	9.0	758	6 Q9N2H7	Q9n2h7 sus scrofa
5	175	8.5	455	11 Q920L8	Q920l8 mus musculus
6	175	8.5	535	11 Q9EQT7	Q9egt7 mus musculus
7	161	7.8	307	11 Q54947	Q54947 rattus norv
8	143.5	7.0	299	4 Q9UBK4	Q9ubk4 homo sapien
9	143.5	7.0	299	4 Q9UGN4	Q9ugn4 homo sapien
10	126	6.1	359	4 Q4356	Q4356 homo sapien
11	124	6.0	288	4 Q9HD97	Q9hd97 homo sapien
12	122.5	6.0	301	4 Q95100	Q95100 homo sapien
13	118	5.7	820	4 Q60585	Q60585 homo sapien
14	117	5.7	897	11 Q70495	Q70495 mus musculus
15	116.5	5.7	364	4 Q96D42	Q96d42 homo sapien
16	114.5	5.6	335	13 Q9YGV5	Q9ygv5 gallus gall

17	113.5	5.5	1417	12	067631	Q67631 gallid herp
18	113	5.5	1537	5	Q9VAI2	Q9vai2 drosophila
19	112	5.5	2082	2	Q9S200	Q9s200 streptomyce
20	111.5	5.4	392	5	Q44716	Q44716 caenorhabdi
21	111	5.4	367	10	Q9AYC9	Q9ayc9 oryza sativ
22	111	5.4	801	5	Q23635	Q23635 caenorhabdi
23	109.5	5.3	335	13	Q9PWR4	Q9pwr4 gallus gall
24	109.5	5.3	862	11	Q9JIK1	Q9jik1 rattus norv
25	109.5	5.3	2321	12	Q9DGT6	Q9dgt6 turkey herp
26	109	5.3	666	11	Q9NVL2	Q9nvl2 mus musculu
27	108.5	5.3	923	4	Q9NVB8	Q9nvb8 homo sapien
28	108	5.3	892	4	Q9Y438	Q9y438 homo sapien
29	107.5	5.2	528	16	Q9RSJ1	Q9rsj1 deinococcus
30	107	5.2	270	4	Q9UMT1	Q9umt1 homo sapien
31	107	5.2	270	4	Q9H564	Q9h564 homo sapien
32	107	5.2	355	2	Q93RI1	Q93rl1 streptococ
33	107	5.2	385	2	Q54913	Q54913 streptococ
34	107	5.2	1275	4	Q9UQ36	Q9uq36 homo sapien
35	107	5.2	1783	4	Q15038	Q15038 homo sapien
36	107	5.2	1791	4	Q60382	Q60382 homo sapien
37	107	5.2	1847	5	Q9KN5	Q9kn5 leishmania
38	107	5.2	2296	4	Q9UHA8	Q9uha8 homo sapien
39	107	5.2	2752	4	Q9UQ35	Q9uq35 homo sapien
40	106.5	5.2	335	13	Q9YGH1	Q9ygh1 gallus gall
41	106.5	5.2	460	10	Q9LFA8	Q9lfa8 arabidopsis
42	106	5.2	390	4	Q96T50	Q96t50 homo sapien
43	106	5.2	390	4	Q96AP7	Q96ap7 homo sapien
44	106	5.2	496	10	Q23094	Q23094 arabidopsis
45	106	5.2	800	4	Q96P24	Q96pz4 homo sapien

ALIGNMENTS

RESULT 1

O60667 PRELIMINARY; PRT; 390 AA.

AC O60667; (TRENBLrel. 07, Created)

DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)

DT 01-AUG-1998 (TRENBLrel. 19, Last annotation update)

DE ANTI-FAS-INDUCED APOPTOSIS (REGULATOR OF FAS-INDUCED APOPTOSIS).

GN TOSO.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98246048; PubMed=9586636;

RA Hitoshi Y., Lorens J., Kitada S.I., Fisher J., LaBarge M., Ring H.Z.,

RA Francke U., Reed J.C., Kinoshita S., Nolan G.P.;

RT "Toso, a cell surface, specific regulator of Fas-induced apoptosis in

RL Immunity 8:461-471(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=PRIMARY B-CELLS FROM TONSILS;

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF057557; AAC18830.1; -.

DR EMBL; BC006401; AAH06401.1; -.

DR InterPro; IPR003599; Ig.

DR InterPro; IPR003006; Ig_MHC.

DR Pfam; PF00047; ig; 1.

DR SMART; SM00409; Ig; 1.

SQ SEQUENCE 390 AA; 43146 MW; FE91D217EBCA99C6 CRC64;

Query Match 99.6%; Score 2047; DB 4; Length 390;
Best Local Similarity 99.7%; Pred. No. 2e-174;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLWPLYFLPVSGALRILPEVKVEGELGSGSVTIKCPLEPMHVRIYLCREMAGSGTCCT 60

```

Db 1 MDFWLMPLYFLPVSGALRILPEVKVEGELGGSVTIKCPLEMHVRIYLCREMGSGTCT 60
QY 61 VSTTNFKAIEYKGRVTLKQYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
Db 61 VSTTNFKAIEYKGRVTLKQYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
QY 121 NVHSEYPSWEQPMPTPKWFLPFLFOMPAYASSSKFVTRVTPAQRGKVPVHSSP 180
Db 121 NVHSEYPSWEQPMPTPKWFLPFLFOMPAYASSSKFVTRVTPAQRGKVPVHSSP 180
QY 181 TQIOTHRPRVSRASSVAGDKPRTFLPSTTASIKSALGGLKQPTSYNHHTRHLRORALD 240
Db 181 TQIOTHRPRVSRASSVAGDKPRTFLPSTTASIKSALGGLKQPTSYNHHTRHLRORALD 240
QY 241 YGSGRGEGQGHILPTIILGLFLALLGLVYKRAVERKALSRRLAVRMALESSQ 300
Db 241 YGSGRGEGQGHILPTIILGLFLALLGLVYKRAVERKALSRRLAVRMALESSQ 300
QY 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPLPAPLOVSESPWLHAPSL 360
Db 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPLPAPLOVSESPWLHAPSL 360
QY 361 KTSCEVSVLYHQPAMMEDSDSDYINVPA 390
Db 361 KTSCEVSVLYHQPAMMEDSDSDYINVPA 390

```

RESULT 2

```

Q9D8T1 ID Q9D8T1 PRELIMINARY; PRT; 422 AA.
AC Q9D8T1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1810037B03RIK PROTEIN.
GN 1810037B03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher M., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
RA Hayashizaki Y.;
RT Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
DR EMBL; AK007714; BAB25207.1;
DR MGI; MGI:1916419; 1810037B05Rik.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00410; IG_Like; 1.
DR SEQUENCE 422 AA; 47532 MW; 2597083A50AD8E6E CRC64;

```

```

Query Match 56.3%; Score 1157; DB 11; Length 422;
Best Local Similarity 57.5%; Pred. No. 5.8e-95;
Matches 234; Conservative 46; Mismatches 107; Indels 20; Gaps 6;

QY 1 MDRWLMPLYFLPVSGALRILPEVKVEGELGGSVTIKCPLEMHVRIYLCREMGSGTCT 60
Db 1 MDRWLMPLYFLPVSGALRILPEVKVEGELGGSVTIKCPLEMHVRIYLCREMGSGTCT 60
QY 61 VSTTNFKAIEYKGRVTLKQYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
Db 61 VSTTNFKAIEYKGRVTLKQYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
QY 121 NVHSEYPSWEQPMPTPKWFLPFLFOMPAYASSSKFVTRVTPAQRGKVPVHSSP 174
Db 121 NVHSEYPSWEQPMPTPKWFLPFLFOMPAYASSSKFVTRVTPAQRGKVPVHSSP 174
QY 175 VHSSTPTQIOTHRPRVSRASSVAGDKPRTFLPSTTASIKSALGGLKQPTSYNHHTRHL 234
Db 175 VHSSTPTQIOTHRPRVSRASSVAGDKPRTFLPSTTASIKSALGGLKQPTSYNHHTRHL 234
QY 181 TQIOTHRPRVSRASSVAGDKPRTFLPSTTASIKSALGGLKQPTSYNHHTRHLRORALD 240
Db 181 TQIOTHRPRVSRASSVAGDKPRTFLPSTTASIKSALGGLKQPTSYNHHTRHLRORALD 240
QY 241 YGSGRGEGQGHILPTIILGLFLALLGLVYKRAVERKALSRRLAVRMALESSQ 300
Db 241 YGSGRGEGQGHILPTIILGLFLALLGLVYKRAVERKALSRRLAVRMALESSQ 300
QY 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPLPAPLOVSESPWLHAPSL 360
Db 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPLPAPLOVSESPWLHAPSL 360
QY 361 KTSCEVSVLYHQPAMMEDSDSDYINVPA 390
Db 361 KTSCEVSVLYHQPAMMEDSDSDYINVPA 390

```

RESULT 3

```

Q96SA2 ID Q96SA2 PRELIMINARY; PRT; 534 AA.
AC Q96SA2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FKSG87 PROTEIN.
GN FKSG87.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang Y.-G., Gong L.;
RT Molecular cloning and characterization of FKSG87, a novel gene
RT located on human chromosome 1.*;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF34295; AAK39522.1;
SQ SEQUENCE 534 AA; 56748 MW; 6EF8050E412AF91C CRC64;

Query Match 9.1%; Score 186; DB 4; Length 534;
Best Local Similarity 22.2%; Pred. No. 3.5e-08;
Matches 98; Conservative 57; Mismatches 151; Indels 136; Gaps 20;

QY 3 RMLWLMPLYFLPVSGALRILPEVK-----VEGELGGSVTIKCP 38
Db 45 RMLWE-GSLFTRTHLRAMTGLRTPSSPLCWRESSEFAAPNSLGRSLVSGEPGAVTIQCH 103
QY 39 LP-----EMHVRIYLCREMGSGTCTGVSTTNFKAIEYKGRVTLKQYPRKNLFLVEVTOL 94
Db 104 YAPSVNRRQRYKWCRLGPPRWICQIVSTNYTHRRYDRVALTFPPGRLVVRUSQL 163
QY 95 TESDSGVYACGAGMNTDRGKTOKVTLNVHSEYPSWEQPMPTPKWFLPFLFOMPAY 153
Db 164 SPDDICGICGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIGIG 213
QY 154 ASSSKFVTRVTPAQRGKVPVHSSPTQIOTHRPRVSRASSVAGDK--PRTFLPSTTAS 211

```



```
Db 214 GTSAPVANRWTP-----GTTQTLOGTANDTVASTPCTSKTTASAEGRTPGATREAPPGT 269
Qy 212 KISALEGLLK-----POTPSYNHHTLRHQRALDYGSQSGREGOGFHLIPTILGLFLLA 266
Db 270 G-SWAGSVKAPAPIPEPPSPKSRMSNTTICGVMEGRSS----- 308
Qy 267 LGLLVKRA--VERKALSRARR-----LAVRMRALESSOR----- 301
Db 309 ----VTNRARASKDRREMTTKADRPREDIEGVRI-ALDAAKKVLGTGPPALVSETLAW 363
Qy 302 ---PRGSP--RPRSONNIYSACPRRARGADAAGTGEAPV----- 335
Db 364 EILQATPVSKQSGSIGETTP--AAGWTLGTTPAADVWITSMEASGESGAAGDLDA 421
Qy 336 ----PGGAPLPPAPLOVSESPW 354
Db 422 TGDGRGPQATLSQTP---AVGPW 440

RESULT 4
Q9N2H7 PRELIMINARY; PRT; 758 AA.
AC Q9N2H7;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE POLY-IG RECEPTOR PRECURSOR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP Sone T., Kumura H.;
RT "Porcine mammary gland cDNA clone, similar to poly-Ig receptor.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB032195; BAA84283.2; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 3.
DR SMART; SM00410; IG_Like; 2.
KW Signal; Receptor.
FT SIGNAL 1 18 POTENTIAL.
SQ SEQUENCE 758 AA; 83154 MW; D5BEBIA8B082D247 CRC64;

Query Match 9.0%; Score 184; DB 6; Length 758;
Best Local Similarity 34.2%; Pred. No. 8.3e-08;
Matches 41; Conservative 22; Mismatches 41; Indels 16; Gaps 4;

Qy 17 LRILPEVKGEL-----GGSVTTKCLP-----EMHVRIYLCREMAGSGTCGTVV 62
Db 10 LAIFPVVSMKSPFGPDVSSVSGSSVIRCYYPATSVNRHSGKYWCR-IGAKGRCTILI 68
Qy 63 STTNFIKAEYKGRVTLKQYPRKNLFLVEVTVLTQTESDGVYACGAGMNTDRGKTQKVTLANV 122
Db 69 SSEGYSIDKYGRANLUNTFNPFNGTFFVNDIGHLTRGDSGLYKCGLGISS-RGLSFDVSLEV 127

RESULT 5
Q920L8 PRELIMINARY; PRT; 455 AA.
AC Q920L8;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE FCA/M RECEPTOR (FRAGMENT).
GN FCAMR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu Y., Honda S., Yotsumoto K., Tahara-Hanaoka S., Eyre H.J.,
RA Sutherland G.D., Endo Y., Shibuya K., Koyama A., Nakauchi H.,
RA Shibuya A.;
RT "Fca/m receptor is a single gene-family member closely related to
RT polymeric immunoglobulin receptor on chromosome 1.";
RL Immunogenetics 0:0-0(2001).
DR EMBL; AB071978; BAB71750.1; -.
KW Receptor.
FT NON_TER 1
SQ SEQUENCE 455 AA; 48810 MW; 628913C33A4AC365 CRC64;

Query Match 8.5%; Score 175; DB 11; Length 455;
Best Local Similarity 23.5%; Pred. No. 2.7e-07;
Matches 89; Conservative 39; Mismatches 136; Indels 114; Gaps 14;

Qy 25 VEGELGSGVTTKCLP-----EMHVRIYLCREMAGSGTCGTVVSTTNFIKAEYKGRVTLKQ 80
Db 10 VTGNTGGAVTTHCHYAPSSVNRHQRKYWCRLGSLPWICHTVWSTNQYTHPDYRGRAALTD 69
Qy 81 YPRKNLFLVEVTVLTQTESDGVYACGAGMNTDRGKTQKVTLANVHSEYEPS---WEEQPMPE 137
Db 70 VPQSGLFVVRLLRLSLGDLVGLYRCGIG---DRNDMLFVSFVNLTVSAGPSNTTYAAAPASS 126
Qy 138 TPKWFHLPYLPQMPAYASSSKFVTRVTTPAQGRKVPVPHHSSPTTQIHRPRVSRASSVA 197
Db 127 EP-----TTASPCAASS-----ACNG-----WTSQVTVILEG-----S 154
Qy 198 GDKPRTFLPSTASKISALEGLLKQPQPSYNNHHTLRHQRALDYGSQSGREGOGFHLIP 257
Db 155 GSEWDRTPATTGTSK-----TTSSANGROTLTARTVTLGTGSREES----- 197
Qy 258 TILGLFLLALLGLVYKRAVERRKALSRARLAVRMR-----ALESSQRPGR 304
Db 198 -----IRAAVPTPEGPSPKRSMSSTTQCVLWNTNSVTPSVTTSEGRQ 243
Qy 305 SPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPLPAPL-----QVSE 351
Db 244 GTTPETDG-----PRDE--TDVRSPEAPRKTGTTTPSALISEHVHTWETLQDKTEVSK 295
Qy 352 SPWLH-----APSLKT 362
Db 296 QOMLHSLLEELSPAPSAQT 313

RESULT 6
Q9EQT7 PRELIMINARY; PRT; 535 AA.
AC Q9EQT7;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE FCA/M RECEPTOR.
GN FCAMR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21170225; PubMed=11062505;
RA Shibuya A., Sakamoto N., Shimizu Y., Shibuya K., Osawa M.,
RA Hiroyama T., Eyre H.J., Sutherland G.R., Endo Y., Fujita T.,
RA Miyabayashi T., Sakano S., Tsuji T., Nakayama E., Phillips J.H.,
RA Lanier L.L., Nakauchi H.;
RT "Fca/m receptor mediates endocytosis of IgM-coated microbe.";
RL Nat. Immunol. 1:441-446(2000).
DR EMBL; AB048834; BAB17312.1; -.
DR MGD; MGI:1927803; Fcamr.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
```

```
DR InterPro: IPR000508; Peptidase_S26.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; IG; 1.
DR SMART: SM00410; IG.Like; 1.
DR PROSITE: PS00501; SPASE_1_1; UNKNOWN_1.
DR Receptor.
SQ SEQUENCE 535 AA; 57696 MW; B275B77C70151C75 CRC64;

Query Match
Best Local Similarity 23.5%; Score 175; DB 11; Length 535;
Matches 89; Conservative 39; Mismatches 136; Indels 114; Gaps 14;

QY 25 VEGELGGSVTIKCLPLP-----EMHVIYLCREMAGSGTCGTVVSTNTHIKAEYKGRVTLKQ 80
DB 90 VTGNTGGAVTHCHYAPSSVNRHQKRWCRGLGSLPWLWCHTIVVSTNQTHDPYDGRALTD 149
QY 81 YPKNLFVLEVTQLTSDSGVYACAGNMTDRGKTQKVTLVNHSEYEPS---WEEQPMPE 137
DB 150 VPQSLGFVYVRLRLSLDGLVYRCIG---DRNDMLFFSVNLTVSAGPSNTTAAAPASS 206
QY 138 TPKWHLPLYPFQMPAYASSSKFVTRVTPPAQRGKVPVPHHSPTQTTHRPRVSRASSVA 197
DB 207 EP-----TTASPGAASS-----ACNG-----WTSQVTLLEG-----S 234
QY 198 GDKPTFLPSTTASKISALEGLLKQPOTPSYNNHTRLHQRALDYGSGSGREGGQGFILIP 257
DB 235 GSEWDRTPATPTGTSK-----TTSSANGROTTLRTARTVVLGTGSRREGS----- 277
QY 258 TILGLFLLALGLVWKRVAERKALSRARRLAVRMK-----ALESSORPRG 304
DB 278 -----IRAAVPTPEGSPKSRSSSTTQGVWLNRNVTSPVTSVTTSEGRQ 323
QY 305 SPRPSQNNIYACPRRAGADAAGTGAEPVPGAPLPPAPL-----QVSE 351
DB 324 GTTPETDG-----PRDE--TDVRVSPAPRKTGTGTRPSALISEHVHTWETLQDKTEVSK 375
QY 352 SPWLH-----APSLKT 362
DB 376 QQLHLSLEELSPAPSAQT 393

RESULT 7
05497 ID 054947 PRELIMINARY; PRT; 307 AA.
AC 054947;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DE KIDNEY INJURY MOLECULE-1 PRECURSOR (KIM-1).
GN KIM-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=98129827; PubMed=9461608;
RA Ichimura T., Bonventre J.V., Bailly V., Wei H., Hession C.A.,
RA Cate R.L., Sanicola M.;
RT "Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion
RT molecule containing a novel immunoglobulin domain, is up-regulated in
RT renal cells after injury.";
RL J. Biol. Chem. 273:4135-4142(1998).
CC -!- FUNCTION: MAY ACT AS AN EPITHELIAL CELL ADHESION MOLECULE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN LIVER, SPLEEN AND
CC -!- EXPRESSION FOUND IN REGENERATING PROXIMAL TUBULE EPITHELIAL CELLS.
CC -!- INDUCTION: IN RENAL CELLS, AFTER INJURY.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC ONE V-LIKE DOMAIN.
DR EMBL; AF035963; AAC53546.1; -.
```

```
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; IG; 1.
DR SMART: SM00410; IG.Like; 1.
DR PROSITE: PS00501; SPASE_1_1; UNKNOWN_1.
DR Receptor.
SQ SEQUENCE 535 AA; 57696 MW; B275B77C70151C75 CRC64;

Query Match
Best Local Similarity 25.1%; Score 161; DB 11; Length 307;
Matches 89; Conservative 52; Mismatches 119; Indels 94; Gaps 20;

QY 13 VSGALRLIP-----EVKVEGELGGSVTIKCLPLPMHVRIYLC--REMAGSGTCGVVST 64
DB 8 ISGLLLLLGVSDVSEV-VKGVGHVPVTPCTYSTRGGITTCWGRGQCPSYSCQNLILW 66
QY 65 TNFIKAEYK--GRVTLKQYPRKNLFVLEVTQLTSDSGVYACGA--GMNTRGRKTQVT 119
DB 67 TNGVTVYTSRGRYNIKRISEGDVSLTIENSVDLSGLYCCRVETPGWFDN---QKMT 122
QY 120 LNVHSEYFSEWEQPMPEPKWHLPLYPFQMPAYASSSKFVTRVTPPAQRGKVPV---V 175
DB 123 FSL--EVKPEITSP--PTRP-----TTTPRTTPTTISTRTHVPTSTRV 165
QY 176 HHSPT---TQTHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQPOTPSY---- 227
DB 166 STSTPTPEQTQ--THRPEIT-----TFYAHETAEVT-----ETPSYTPADW 205
QY 228 -----NHETRLHQRALDYGSGSGREGGQGFILIPILGLFLLALLGLVWKR 275
DB 206 NGVTVSECAWNNHTVRPLRK-----PQRNPTKGFYGVGMSVAALLLLLASTVWVTRY 259
QY 276 VERRK---ALSRARRLAVRMRALESSORPRGSPRSONNIYACPRRAGAD 326
DB 260 IIRKKMGLSLFVAFHVS-KSRALQNA----AIVHPRAEDNIY-IIEDSRGAE 307

RESULT 8
09UBK4 ID 09UBK4 PRELIMINARY; PRT; 299 AA.
AC 09UBK4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE IRL1.
GN IRL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOID;
RA O'Connor C.D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOID;
```


Db 249 PPREEVEYST---VASPREELHYASVFDSTNTNRIAAQRPREEPDS 294

RESULT 13

060585 PRELIMINARY; PRT; 820 AA.

AC O60585;

DT 01-AUG-1998 (TEMBLrel. 07, Created)

DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)

DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)

DE SER/ARG-RELATED NUCLEAR MATRIX PROTEIN.

GN SRM160.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Blencowe B.J., Issner R., Nickerson J.A., Sharp P.A.;

RT "A Coactivator of Pre-mRNA Splicing.";

RL Genes Dev. 0:0-0(1998).

DR EMBL: AF048977; AAC09321.1; -.

DR InterPro: IPR002483; PWI.

DR Pfam: PF01480; PWI; 1.

DR SMART: SM00311; PWI; 1.

KW Matrix protein.

SO SEQUENCE 820 AA; 93519 MW; B82F76DC6108D750 CRC64;

Query Match 5.7%; Score 118; DB 4; Length 820;

Best Local Similarity 21.8%; Pred. No. 0.071;

Matches 66; Conservative 33; Mismatches 102; Indels 102; Gaps 11;

QY 112 RGKTKVTLNVHSEYEPSEQPMPTPKWFLPYLFQMPAYASSKVFTR-----VTPP 166

Db 346 RRRSSASLGSSSSSSSSRSPKPP-----PKRTSSPPKTRRLSPSASP 394

QY 167 AQRGK-----VPPVHSSPTQITHRPRVSRASSVAGDKPRTFLPSTTASIKISALEGLK 221

Db 395 RRRHRPSPATPPKTRDSPTPQOSNTRKSRVS-----VSPGRTSGKVTXKKGTEK 445

QY 222 PQTPS-----YNHHTRLHQRALDYGSQSGREGQG 252

Db 446 RESPSAPKPRKVELSESEEDKGGKMAADSVQORRQYRRQNOQSSSDSGSSSTSEDER 504

QY 253 HILPTILGLFLLALLGLVVKRAVERRKALSRARRLAVRM-----RALESSQRPRG--- 304

Db 505 -----PKRSHVKNGEVGRRRRSPSPKRRQKETSPPRGR 545

QY 305 -SPRP-----RSQNNIYACPRRARGADAAGTGEAPVPGGAPLP-PAPLQVSESPWLHAP 358

Db 546 RSPSPPTRRRRSPSPAPPPRRR-----TPTPPRRRTSPSPPPRRRSPSPRYS 596

QY 359 SLK 361

Db 597 PIQ 599

RESULT 14

070495 PRELIMINARY; PRT; 897 AA.

AC O70495;

DT 01-AUG-1998 (TEMBLrel. 07, Created)

DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)

DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)

DE PLENTY-OF-PROLINES-101.

GN SRM1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RA Vayssiere B.M., Camonis J.H.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF062655; AAC17422.1; -.

DR MGD: MGI:1858303; Srrml.

DR InterPro: IPR002483; PWI.

DR Pfam: PF01480; PWI; 1.

DR SMART: SM00311; PWI; 1.

SO SEQUENCE 897 AA; 101166 MW; 621608EB6772BD10 CRC64;

Query Match 5.7%; Score 117; DB 11; Length 897;

Best Local Similarity 21.5%; Pred. No. 0.098;

Matches 62; Conservative 31; Mismatches 108; Indels 88; Gaps 7;

QY 112 RGKTKVTLNVHSEYEPSEQPMPTPKWFLPYLFQMPAYASSKVFTRVTT----- 165

Db 345 RRRSSASLGSSSSSSSSRSPKPP-----PKRTSSPPKTRRLSPSASP 393

QY 166 -----PAORGKVPVHSSPTQITHRPRVSRASSVAGDKPRTFLPSTTASIKISALEGLL 220

Db 394 RRRHRPSPATPPKTRHSPTPQOSNTRKSRVS-----VSPGRTSGKVTXKKGTE 444

QY 221 KPQTPS-----YNHHTRLHQRALDYGSQSGREGQG 251

Db 445 KRESPPAPKPRKVELSESEEDKGGKMAADSVQORRQYRRQNOQSSSDSGSSSTSEDER 504

QY 252 PHILPTILGLFLLALLGLVVKRAVERRKALSRARRLAVRMRALESSORPGRSPRSQ 311

Db 505 -----PKRSHVKNGEVGRRRRLSPSRASPSPRKQKETSPPRMQ 543

QY 312 NN-----IYACPRRARGADAAGTGEAPVPGGAPLPAPLQVSESP 353

Db 544 MGRMQSPVTKSSRRRSPSPPPARRRRSPSPAPPPPPPPPPRRRRSP 592

RESULT 15

096D42 PRELIMINARY; PRT; 364 AA.

ID Q96D42;

AC Q96D42;

DT 01-DEC-2001 (TEMBLrel. 19, Created)

DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 39.2 KDA PROTEIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=RENAL ADENOCARCINOMA;

RA Strausberg R.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC013325; AAH13325.1; -.

KW Hypothetical protein.

SO SEQUENCE 364 AA; 39249 MW; AA67C7DC7FAC81F1 CRC64;

Query Match 5.7%; Score 116.5; DB 4; Length 364;

Best Local Similarity 20.5%; Pred. No. 0.034;

Matches 73; Conservative 54; Mismatches 144; Indels 85; Gaps 13;

QY 23 VKVEGLGSGVTIKCPLEPMHVIYLCREMGAGSGTCCTVVTNFIKAEYKGRVTLK--- 79

Db 22 VKVGEAGPSVTLPCYSGAVTSMCNWRGSCSLFTQCQNGIVWTNGTHVTYRKDTRYKLLG 81

QY 80 QYPRKNFLVEVTQLTESDSGVYACGAGMTDRCKTKVTLNVHSEYEPSEQPMPTP 139

Db 82 DLSSRDVSLT-IENTAVSDSGVYCCRV---EHRGWFMNMKITVSLEIVP-----P 127

QY 140 KWFHLPYLFQMPAYA---SSSKFVTRVTP---AQRGKVPVPHHSSPTQITHRPRVSR 193

Db 128 KYTTTPITVTVPTVTVTSTVTTTTPMTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 187

QY 194 SSVAGDKPRTFLPSTTASIKI-----SALSG 218

Db 188 TSV---PTTSIPTTSVPVTTTSTFVPMPPLPRONHEPVATSPSPQPAETHETTLQG 244
Qy 219 LKPO---TPSYNHHT-----RLHRQALDYGSQSGREGQGFHILPTIL 260
Db 245 AIRREPTSSPLYSYTTDGDNDVTESDGLWNNNOTQLFLEHSLLTANTTKGIYAGV-CIS 303
Qy 261 GLFLLALLGLVVKRAVERRKALSRARRL-AVRMRALESSORPRGSPRPRSONNIY 315
Db 304 VLVLALLGLVIAKKYFFKKEVQQLSVFSSLOIKALONAV-----EKEVQAEADNIY 355

Search completed: October 28, 2002, 17:32:33
Job time : 39.121 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:29:31 ; Search time 12.371 Seconds
(without alignments)
1833.089 Million cell updates/sec

Title: US-09-135-238B-2_COPY_18_253
Perfect score: 1254
Sequence: 1 RILPEVKVEGLGGVTIK.....HRQRALDYGSGREGQGPFH 236

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	14.8	769	1 ORRTGS	secretory componen
2	180.5	14.4	764	1 ORHUGS	secretory componen
3	179.5	14.3	757	1 S48841	secretory componen
4	179	14.3	757	2 I45956	polymERIC immunogl
5	166.5	13.3	773	1 QRRBG	secretory componen
6	100.5	8.0	504	2 A56205	transcription fact
7	100	8.0	224	2 I37243	CMRP-35 antigen -
8	96.5	7.7	847	2 JH0371	B-cell adhesion pr
9	93	7.4	346	2 S17475	STES0 protein - ye
10	92.5	7.4	504	2 I4257	NF2d9 - mouse
11	92	7.3	355	1 LKCH	proteoglycan link
12	92	7.3	1906	1 S68235	myosin-light-chain
13	91.5	7.3	502	2 C56205	transcription fact
14	90.5	7.2	534	2 T2154	hypothetical prote
15	90.5	7.2	556	2 I45066	steroid hormone re
16	90.5	7.2	1839	1 RRPPEM	genome polyprotein
17	89.5	7.1	1415	1 EDBEGA	immediate-early pr
18	89	7.1	1241	2 S01827	period clock prote
19	88.5	7.1	328	2 JQ0985	hydroxyproline-ric
20	88.5	7.1	1021	2 T42634	connectin/titin -
21	87	6.9	249	2 S69340	Ig heavy chain VHI
22	87	6.9	1070	2 T34385	hypothetical prote
23	86.5	6.9	228	2 JC7761	dendritic cell-der
24	86.5	6.9	456	2 T45610	proanthranilate N-
25	86	6.9	303	2 S28264	hydroxyproline-ric
26	86	6.9	2453	2 S60254	nuclear receptor c
27	85.5	6.8	342	2 A46529	Ig gamma chain (5)
28	85.5	6.8	496	2 T01564	hypothetical prote
29	85.5	6.8	501	2 A42030	alpha-globin trans

30	85.5	6.8	1506	2 JC5985	phosphoinositide 3
31	85.5	6.8	3649	1 S18268	delta-(L-alpha-ami
32	85	6.8	283	2 S13383	hydroxyproline-ric
33	85	6.8	350	2 S22456	hydroxyproline-ric
34	85	6.8	2697	2 T25444	hypothetical prote
35	84.5	6.7	594	2 S50611	hypothetical prote
36	84.5	6.7	704	1 S33704	transforming prote
37	84	6.7	275	2 A45679	inhibitor-of-apopt
38	84	6.7	363	2 A72702	hypothetical prote
39	84	6.7	369	2 S20500	hydroxyproline-ric
40	84	6.7	460	2 T45968	hypothetical prote
41	84	6.7	467	1 HLMSP3	poliovirus recepto
42	84	6.7	530	2 A53437	poliovirus recepto
43	84	6.7	850	2 S56015	gastric mucin MUC5
44	83.5	6.7	137	2 S78054	Ig heavy chain pre
45	83.5	6.7	146	1 G1HUH2	Ig heavy chain pre

ALIGNMENTS

RESULT 1

ORRTGS

secretory component precursor - rat

N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor

N:Contains: free secretory component; transmembrane secretory component

C:Species: Rattus norvegicus (Norway rat)

C:Date: 07-Sep-1990 #sequence_revision 23-Aug-1996 #text_change 16-Jul-1999

C:Accession: S05407; S54731

R:Banting, G.; Brake, B.; Braghetta, P.; Luzio, J.P.; Stanley, K.K.

FEBS Lett. 254, 177-183, 1989

A:Title: Intracellular targeting signals of polymeric immunoglobulin receptors are h

A:Reference number: S05407; MUID:89378226

A:Accession: S05407

A:Molecule type: mRNA

A:Residues: 1-769 <BAN>

A:Cross-references: EMBL:X15741; NID:g56464; PIDN:CAA33758.1; PID:g56465

C:Complex: monomeric as a transmembrane receptor or free in mucosal secretions; heter

amers: hetero-22-mer composed of one chain of secretory component, one chain of immun

C:Superfamily: secretory component; glycoprotein; immunoglobulin homology

C:Keywords: duplication; signal sequence #status predicted <SIG>

F:1-18/Domain: immunoglobulin homology

F:19-769/Product: transmembrane secretory component #status predicted <MATW>

F:19-582/Product: free secretory component #status predicted <MATF>

F:33-112/Domain: immunoglobulin homology <IMI>

F:145-222/Domain: immunoglobulin homology <IM2>

F:250-326/Domain: immunoglobulin homology <IM3>

F:363-442/Domain: immunoglobulin homology <IM4>

F:477-548/Domain: immunoglobulin homology <IM5>

F:644-666/Domain: transmembrane #status predicted <TMM>

F:667-769/Domain: intracellular #status predicted <INT>

F:40-110/56-64/152-220/370-440/384-394/484-546/498-505/Disulfide bonds: #status predi

F:90/135/471/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:488-522/Disulfide bonds: (in Ig-unbound form) #status predicted

F:522/Binding site: interchain (to IgA alpha-1 chain-192) #status predicted

F:678/Binding site: cysteine (Cys) (covalent) (in Ig-bound form) #status predicted

F:678/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 14.8%; Score 185; DB 1; Length 769;

Best Local Similarity 34.5%; Pred. NO. 4e-07;

Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 13 GGSVTKIKPLPE----MHVRIYLCEMAGSGTCGTVSTTFPIKAEYKRVTLKQYPRKN 68

Db 33 GMSVITCYTPDTSVNRHTRKYWCQGA-NGYCATLISSNGYLSKSGRASLIINFPNS 91

QY 69 LFLVETVQLTSDSGVYACGAGMNTDRGKTOKVTLNHSEYEPSEWEEQPMPTKWFHL 127

Db 92 TPVINIAHLTQEDTGSYKCGLG-TTNRLGLFEDVSLV-----SQVFEFPNDTHV 139

RESULT 2

QRHUGS

QY 4 PEVKVEGELGSGVTIKCPPLP-----EMHVRIVLCREMAGSGTCGTVVSTTNFIKAEYKGRV 59
Db 25 PE-EVSSVEGRSVSIKYYPTTSVNRHTRKYWCROGA-QGRCTTLLISSEGIVSDDYVGRA 82
QY 60 TLQYPRKNLFLVEVTLQTESDSGVACGAGMNTDRGKTOKVTLNVHSE 108
Db 83 NLTNFPESGTFVVDISHLTHKDSGRYKGLGISS-RGLNFDVSLEVSQD 130
RESULT 4
I45956
polymeric immunoglobulin receptor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 23-Jul-1999
R:Kulseth, M.A.; Krajci, P.; Myklebost, O.; Rogne, S.
DNA Cell Biol. 14, 251-256, 1995
A:Title: Cloning and characterization of two forms of bovine polymeric immunoglobulin re
C:Accession: I45956
A:Reference number: I45956; MUID:95186063
A:Accession: I45956
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-757 <KUL>
A:Cross-references: GB:L04797; NID:g388279; PIDN:AAC41620.1; PID:g388280
C:Superfamily: secretory component; immunoglobulin homology
F:145-222/Domain: immunoglobulin homology <IMM>
Query Match 14.3%; Score 179; DB 2; Length 757;
Best Local Similarity 33.9%; Pred. No. 1.2e-06;
Matches 41; Conservative 21; Mismatches 43; Indels 16; Gaps 4;
QY 2 ILPEVKVEGEL-----GGSVTIKCPPLP-----EMHVRIVLCREMAGSGTCGTVVST 47
Db 12 IFPVVMSKPIFGPEEVTSEGRSVSIKYYPTTSVNRHTRKYWCROGA-QGRCTTLLIS 70
QY 48 TNFIKAEYKGRVTLQYPRKNLFLVEVTLQTESDSGVYACGAGMNTDRGKTOKVTLNVHS 107
Db 71 EGYVSDDYVGRANLTNPFESGTFVVDISHLTHKDSGRYKGLGISS-RGLNFDVSLEVSQ 129
QY 108 E 108
Db 130 D 130
RESULT 5
QRRBG
secretory component precursor - rabbit
N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
N:Contains: free secretory component; transmembrane secretory component
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
C:Accession: A02111; A28077
R:Mostov, K.E.; Friedlander, M.; Blobel, G.
Nature 308, 37-43, 1984
A:Title: The receptor for transepithelial transport of IgA and IgM contains multiple imm
A:Reference number: A02111; MUID:8414246
A:Accession: A02111
A:Molecule type: mRNA
A:Residues: 1-773 <MOS>
A:Cross-references: GB:X00412; GB:K01291; NID:g1595; PIDN:CAA25118.1; PID:g1596
A:Note: The authors translated the codon ACC for residue 54 as Asn
R:Frutiger, S.; Hughes, G.J.; Hanly, W.C.; Jaton, J.C.
J. Biol. Chem. 263, 8120-8125, 1988
A:Title: Rabbit secretory components of different allotypes vary in their carbohydrate c
A:Reference number: A28077; MUID:88228032
A:Accession: A28077
A:Molecule type: protein
A:Residues: 87-114;410-424 <FRU>
C:Comment: This receptor binds polymeric IgA and IgM at the basolateral surface of epith
process, cleavage occurs to separate the extracellular portion, also known as the secret
C:Comment: The five domains exhibit homology with immunoglobulin V regions. The similar
C:Comment: Alternative splicing in the extracellular domain leads to high or low molecu

C:Superfamily: secretory component; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; immunoglobulin receptor;
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-773/Product: transmembrane secretory component #status predicted <MATM>
F:19-575/Product: free secretory component #status predicted <MATF>
F:30-647/Domain: extracellular #status predicted <EXT>
F:39-117/Domain: immunoglobulin homology <IM1>
F:148-227/Domain: immunoglobulin homology <IM2>
F:253-326/Domain: immunoglobulin homology <IM3>
F:362-440/Domain: immunoglobulin homology <IM4>
F:471-540/Domain: immunoglobulin homology <IM5>
F:648-670/Domain: transmembrane #status predicted <TMM>
F:671-773/Domain: intracellular #status predicted <INT>
F:648-115,155-225,260-324,369-438,478-538/Bisulfide bonds: #status predicted
F:108/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F:418/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 13.3%; Score 166.5; DB 1; Length 773;
Best Local Similarity 42.7%; Pred. No. 1.2e-05;
Matches 44; Conservative 11; Mismatches 41; Indels 7; Gaps 4;
QY 13 GGSVTIKCPPLP-----EMHVRIVLCREMAGSGTCGTVVSTTNFIKAEYKGRVTLKQYPRKN 68
Db 39 GDSVSTICYPTTSVTRHSRKFMCRE-EESGRCVTLAS-TGYTSOEYSGRGLTDFDPKG 96
QY 69 LFLVEVTLQTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEP 111
Db 97 EFPVTVQDLTQDSSGYSKCGGVN-GRGLDFGVNVLVSQKPEP 138
RESULT 6
A56205
transcription factor LBPla - human
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C:Accession: A56205
R:Yoon, J.B.; Li, G.; Roeder, R.G.
Mol. Cell. Biol. 14, 1776-1785, 1994
A:Title: Characterization of a family of related cellular transcription factors which
A:Reference number: A56205; MUID:94158849
A:Accession: A56205
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-504 <VOO>
C:Keywords: alternative splicing; transcription factor
Query Match 8.0%; Score 100.5; DB 2; Length 504;
Best Local Similarity 25.6%; Pred. No. 1.6;
Matches 50; Conservative 30; Mismatches 68; Indels 47; Gaps 11;
QY 47 TTNFIKAEYKGRVTLKQYPRKN-----LFLVEVTLQTESDSGVY-----AC----- 87
Db 180 TSAFIQVHC---ISTEFTPRKHGGEKGVPRIOVDFKQNGEYTDHLHSASQCIKVKFK 236
QY 88 --GAG--MMNTDRGKTOKVTLNVHSEYEPGWEEQPMPTPKFPHLPYLFOMPAYASSSKVF 143
Db 237 PKGADRKQKTDREKMEKTAHEKYPQSYDTTILTECSPWPDAP-----TAYVNNSPSP 291
QY 144 TRVYTPAQRG-KVPVPHSSPTTQITHPRVSRSSVAGDKPRTFLPSTTASKISALEG 201
Db 292 APTFTSPQSTCVSPDSNSSPNHQ-----GDGASQTSGEQ---IQPSATIQETO--QW 340
QY 202 LLKPQTPSYNNHTRL 216
Db 341 LLKNRPFSSY---TRL 352
RESULT 7
I37243
CMRF-35 antigen - human
C:Species: Homo sapiens (man)
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I37243

R:Jackson, D.G.; Hart, D.N.; Starling, G.; Bell, J.I.
Eur. J. Immunol. 22, 1157-1163, 1992
A:Title: Molecular cloning of a novel member of the immunoglobulin gene superfamily hom
A:Reference number: I37243; MUID:92249405
A:Accession: I37243
A>Status: preliminary: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-224 <RES>
A:Cross-references: EMBL:X66171; NID:g396169; PIDN:CAA46948.1; PID:g396170
C:Genetics:
A:Gene: CMK35

Query Match 8.0%; Score 100; DB 2; Length 224;
Best Local Similarity 31.3%; Pred. No. 0.7;
Matches 26; Conservative 15; Mismatches 38; Indels 4; Gaps 3;

QY 8 VEGEIGGVTIKCPLPMH--VRIYLCREMAGSGTCTGVVSTNFIKAEYKGRVTLKQYP 65
Db 31 VAGPVGGSLSVOCKRYEKEHRTLNKFWCRP-PQILRCDKIVETKG-SACKRNGRVSTROSP 88
QY 66 RKNFLFVEVTOLTESDGSVYACG 88
Db 89 ANLSFTVTLNTEEDAGTYWCG 111

RESULT 8
JH0371
B-cell adhesion protein CD22 beta splice form precursor - human
N:Alternate names: B-cell membrane protein CD22
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 31-Jan-2000
C:Accession: JH0371; I56171
R:Wilson, G.L.; Fox, C.H.; Fauci, A.S.; Kehrl, J.H.
J. Exp. Med. 173, 137-146, 1991
A:Title: cDNA cloning of the B cell membrane protein CD22: a mediator of B-B cell intera
A:Reference number: JH0371; MUID:91086838
A:Accession: JH0371
A:Molecule type: mRNA
A:Residues: 1-847 <WTL>
A:Cross-references: GB:X59350; NID:g36090; PIDN:CAA42006.1; PID:g36091
A:Experimental source: B Lymphocyte
A>Note: the authors translated the codon AAT for residue 358 as Met
R:Wilson, G.L.; Najfeld, V.; Kozlow, E.; Menniger, J.; Ward, D.; Kehrl, J.H.
J. Immunol. 150, 5013-5024, 1993
A:Title: Genomic structure and chromosomal mapping of the human CD22 gene.
A:Reference number: I56171; MUID:93267103
A:Accession: I56171
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 121-269, 'T', 271-473, 'K', 475-614, 'R', 616-638, 'Y', 640-711, 777-847 <WIL2>
A:Cross-references: GB:S61375; NID:g385980; PIDN:AAC18956.1; PID:g3184492
C:Genetics:

A:Gene: GDB:CD22
A:Cross-references: GDB:127545; OMIM:107266
A:Map position: 19q13.1-19q13.1
A:Introns: 138/1; 240/1; 329/1; 417/1; 503/1; 591/1; 679/1; 711/2; 804/3
C:Superfamily: immunoglobulin homology
C:Keywords: alternative splicing; B-cell; cell adhesion; dimer; glycoprotein; phosphoric
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-847/Product: B lymphocyte cell adhesion protein #status predicted <MAT>
F:346-398/Domain: immunoglobulin homology
F:609-661/Domain: immunoglobulin homology <IMM1>
F:688-706/Domain: transmembrane #status predicted <TRA>
F:67,101,112,135,164,231,363,445,448,479,574,634/Binding site: carbohydrate (Asn) (coval
F:764,789/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 7.7%; Score 96.5; DB 2; Length 847;
Best Local Similarity 20.7%; Pred. No. 6.2;
Matches 42; Conservative 31; Mismatches 69; Indels 61; Gaps 8;

QY 1 RILPEVKVEGEGSGVTIKCPLPMHVRVLYLCREMAGSGTCTGVVSTNFIKAEYKGRVT 60
Db 247 KVTPSDAIVRE-GDSVTMTC-----EYSSSNPEYTVSM-----LKDGTS 285

QY 61 LKQYPRKRLFLVEVTOLTESDGSVYACGAGMNTDRGKTQKVTLNHSEYSPSWEE----- 115
Db 286 LK---KQNTFTLNREVTQKDSGYCCQVSDVGPGRSEEVFLQVYAPEPSTVQILHSP 342
QY 116 -----QPMETPKWF-----HLPYLFQMPAYASSKFVTR 145
Db 343 AVESQVEFFLCMSLANPLNTVYTHNGKEMQGRTEEKVHPKTI--LPWHAGTYSQVAEN 400
QY 146 VTTPAQRGKVPVPHHSSPTTQIT 168
Db 401 ILGTGORGPGAELDVQYPPKVT 423
RESULT 9
SL17475
STE50 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YCL032w; protein YCL185
C:Species: Saccharomyces cerevisiae
C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jul-2000
C:Accession: SL17475; S30158; S19360
R:Ramezani Rad, M.; Luetzenkirchen, K.; Xu, G.; Kleinhans, U.; Hollenberg, C.P.
Yeast 7, 533-538, 1991
A:Title: The complete sequence of a 11,953 bp fragment from CIG on chromosome III enc
A:Reference number: SL17471; MUID:91377317
A:Accession: SL17475
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-346 <RAM>
A:Cross-references: EMBL:X59720; NID:g1907116; PIDN:CAA42384.1; PID:g5332
R:Ramezani Rad, M.; Xu, G.; Hollenberg, C.P.
Mol. Gen. Genet. 236, 145-154, 1992
A:Title: STE50, a novel gene required for activation of conjugation at an early step
A:Reference number: S30158; MUID:93156679
A:Accession: S30158
A:Molecule type: DNA
A:Residues: 1-346 <RA2>
A:Cross-references: EMBL:Z11116; NID:g3312; PIDN:CAA77462.1; PID:g3313
R:Hollenberg, C.P.; Kleinhans, U.; Luetzenkirchen, K.; Ramezani Rad, M.; Xu, G.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19350
A:Accession: S19360
A:Molecule type: DNA
A:Residues: 1-346 <HOL>
A:Cross-references: EMBL:X59720; NID:g1907116; PID:e264426; PID:g5332; MIPS:YCL032w
C:Genetics:
A:Gene: SGD:STE50
A:Cross-references: SGD:S0000537; MIPS:YCL032w
A:Map position: 3L
C:Superfamily: SAM homology
F:32-101/Domain: SAM homology <SAM>

Query Match 7.4%; Score 93; DB 2; Length 346;
Best Local Similarity 22.2%; Pred. No. 4.2;
Matches 56; Conservative 43; Mismatches 101; Indels 52; Gaps 10;
QY 2 ILPEVKV-----EGELGGSVTIKCPLEPMHVRVLYLCREMAGSGTCGTIV-----STNFF 50
Db 69 LLPELCIQDCQDLCDGDLNKAIFKILINKRDSKLEWKDKQEDMITVLNLYTTTSA 128
QY 51 IKARYGRVTLKQYPRKRLFLVEVTOLTESDGSVYACG-----AGMNTDRGKTQKVTLN 104
Db 129 KLQEFQ-----SQYTLRMDVLDVDMKTSSTSSSPINTHGVSITVPSSNNIIPSSDGYSL 183
QY 105 -----VISEYEPSEWEPQMPETPKFWHLPLYLFQMPAYASSKFVTRVTTTQAQKGVPP 157
Db 184 QTDYFDVFNROSPSRRESPTV-----VFRQPSLSHSKSL-----HDSKNKVPQ 228
QY 158 VHHSSPTTQITTHRPVRSASVAGDKPRFTLPSTTASIKISALEGLLKPTQPSYNNHTRLH 217
Db 229 I-----STNOSHPSAVSTANI-PCPSPNEALQFLASKEDSCERILKNAMKRNLAODW 282
QY 218 RORAL--DYGSQ 227

Search completed: October 28, 2002, 17:34:02
Job time : 15.371 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 6.66129 Seconds
(without alignments)
1371.778 Million cell updates/sec

Title: US-09-135-238B-2_COPY_18_253

Perfect score: 1254

Sequence: 1 RILPEVKVEGLGSGVTIKC.....HQRALDYGSQSGREGQGFH 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	194	15.5	771	1	P1GR_MOUSE
2	185	14.8	769	1	P1GR_RAT
3	180.5	14.4	764	1	P1GR_HUMAN
4	179	14.3	757	1	P1GR_BOVIN
5	166.5	13.3	773	1	P1GR_RABIT
6	120.5	9.6	102	1	P1GR_PIG
7	100	8.0	224	1	CM35_HUMAN
8	96.5	7.7	847	1	CD22_HUMAN
9	93	7.4	346	1	ST50_YEAST
10	93	7.4	2440	1	NCRL_HUMAN
11	92	7.3	355	1	PLK_CHICK
12	92	7.3	1906	1	KMLS_CHICK
13	91.5	7.3	872	1	FPI_MYTICO
14	90.5	7.2	534	1	ICNRD_CAEEL
15	90.5	7.2	1839	1	POLR_EPMV
16	89.5	7.1	1415	1	ICP4_HSMVG
17	89	7.1	1241	1	PER_DROPS
18	88	7.0	835	1	AXNL_BRARE
19	86	6.9	2453	1	NCRL_MOUSE
20	85.5	6.8	1506	1	PK3G_MOUSE
21	85.5	6.8	3649	1	ACVS_NOCLA
22	85	6.8	283	1	EXTN_SORBI
23	84.5	6.7	704	1	MYB_MOUSE
24	84.5	6.7	739	1	FLOB_YEAST
25	84	6.7	275	1	IAP_GVCP
26	83.5	6.7	530	1	VR2_MOUSE
27	83.5	6.7	146	1	HM2I_HUMAN
28	83.5	6.7	325	1	HM06_CAEEL
29	83.5	6.7	387	1	PIGM_RHOSO
30	83	6.6	5147	1	FAT_DRONE
31	82.5	6.6	358	1	KLFI_MOUSE
32	82	6.5	142	1	HV01_RAT
33	82	6.5	749	1	SM3B_HUMAN

34	81.5	6.5	636	1	KDPG_SCHPO	014019 schizosacch
35	81.5	6.5	912	1	PCGB_BOVIN	Q28062 bos taurus
36	81.5	6.5	1257	1	CAML_HUMAN	P32004 homo sapien
37	81.5	6.5	1259	1	CAML_RAT	Q05695 rattus norv
38	81.5	6.5	1448	1	PK3G_HUMAN	O75747 homo sapien
39	81	6.5	248	1	VGL2_EBV	P03218 epstein-bar
40	81	6.5	862	1	SM42_HUMAN	Q92854 homo sapien
41	80.5	6.4	532	1	EYA2_MOUSE	O08575 mus musculu
42	80.5	6.4	938	1	EBN4_EBV	P03203 epstein-bar
43	80	6.4	518	1	TEX5_HUMAN	Q95993 homo sapien
44	80	6.4	571	1	FLB1_CAMJE	P56964 campylobact
45	80	6.4	688	1	YAQA_SCHPO	Q10109 schizosacch

ALIGNMENTS

RESULT 1
PIGR_MOUSE
ID PIGR_MOUSE STANDARD; PRT; 771 AA.
AC 070570;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Polymetric-immunoglobulin receptor precursor (Poly-IG receptor) (PIGR)
DE [Contains: Secretory component].
GN PIGR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX MEDLINE=95138517; PubMed=7836758;
RA Piskurich J.F., Blanchard M.H., Youngman K.R., France J.A.,
RA Kaetzel C.S.;
RT "Molecular cloning of the mouse polymeric Ig receptor. Functional
RT regions of the molecule are conserved among five mammalian species.";
RL J. Immunol. 154:1735-1747(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=98072444; PubMed=9409786;
RA Martin M.G., Gutierrez E.M., Lam J.T., Li T.W.H., Wang J.;
RT "Genomic cloning and structural analysis of the murine polymeric
RT receptor (pigr) gene and promoter region.";
RL Gene 201:189-197(1997).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Liver;
RA de Groot N., Vollebregt E., Lee S.H., Verbeet M.P., de Boer H.A.;
RT "Molecular cloning and exon-intron organization of the gene encoding
RT the murine polymeric immunoglobulin receptor.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC !- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC !- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DF	01-OCT-1996 (Rel. 34, Last sequence update)	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DT	01-MAR-2002 (Rel. 41, Last annotation update)	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
DE	CMRF35 antigen precursor.	OX	NCBI_TaxID=9606;
GN	CMRF35.	RN	[1]
OS	Homo sapiens (Human).	RP	SEQUENCE FROM N.A. (CD22-BETA).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	RC	TISSUE=Tonsil;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	RX	MEDLINE=91086838; PubMed=1985119;
OX	NCBI_TaxID=9606;	RA	Wilson G.L., Fox C.H., Fauci A.S., Kehrl J.H.;
RN	[1]	RT	"cDNA cloning of the B cell membrane protein CD22: a mediator of B-B
RP	SEQUENCE FROM N.A.	RT	cell interactions";
RX	MEDLINE=92249405; PubMed=1349532;	RL	J. Exp. Med. 173:137-146(1991).
RA	Jackson D.G., Hart D.N.J., Starling G., Bell J.I.;	RN	[2]
RT	"Molecular cloning of a novel member of the immunoglobulin gene	RP	SEQUENCE FROM N.A. (CD22-BETA).
RT	superfamily homologous to the polymorphic immunoglobulin receptor.";	RX	MEDLINE=93267103; PubMed=8496602;
RL	Eur. J. Immunol. 22:1157-1163(1992).	RA	Wilson G.L., Najfeld V., Kozlow E., Menniger J., Ward D.,
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).	RA	Kehrl J.H.;
CC	-1- TISSUE SPECIFICITY: PRESENT ON THE SURFACE OF MONOCYTES,	RT	"Genomic structure and chromosomal mapping of the human CD22 gene.";
CC	NEUTROPHILS, A PROPORTION OF PERIPHERAL BLOOD T AND B LYMPHOCYTES	RL	J. Immunol. 150:5013-5024(1993).
CC	AND LYMPHOCYTIC CELL LINES.	RN	[3]
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.	RP	SEQUENCE FROM N.A. (CD22-BETA).
CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.	RA	Lamerdin J.E., McCreedy P., Adamson A.W., Burkhart-Schultz K.,
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	RA	Garcia E., Kyle A., Ramirez M., Stilwagen S., Garnes J.,
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	RA	Danghan L., Bruce R., Quan G., Montgomery M., Ow D.,
CC	the European Bioinformatics Institute. There are no restrictions on its	RA	Kobayashi A., Olsen A.O., Carrano A.V.;
CC	use by non-profit institutions as long as its content is in no way	RL	Submitted (AUG-1996) to the EMBL/GenBank/DBAJ databases.
CC	modified and this statement is not removed. Usage by and for commercial	RN	[4]
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	RP	SEQUENCE FROM N.A. (CD22-ALPHA).
CC	or send an email to license@isb-sib.ch).	RX	MEDLINE=90231465; PubMed=1691828;
CC	-----	RA	Stamenkovic I., Seed B.;
CC	EMBL; X66171; CAA46948.1; -.	RT	"The B-cell antigen CD22 mediates monocyte and erythrocyte adhesion.";
DR	InterPro: IPR003599; Ig.	RL	Nature 345:74-77(1990).
DR	InterPro: IPR003006; Ig_MHC.	RN	[5]
DR	Pfam: PF00047; Ig 1.	RP	VARIANTS GLU-152; GLY-664; CYS-669 AND ASP-745.
DR	SMART; SM00409; IG; 1.	RX	MEDLINE=99180618; PubMed=10079291;
KW	Antigen; Transmembrane; Glycoprotein; Signal.	RA	Hatta Y., Tsuchiya N., Matsushita M., Shiota M., Hagiwara K.,
FT	SIGNAL 1 20	RA	Tokunaga K.;
FT	CHAIN 21 224	RT	"Identification of the gene variations in human CD22.";
FT	DOMAIN 21 183	RL	Immunogenetics 49:280-286(1999).
FT	TRANSMEM 184 204	CC	-1- FUNCTION: MEDIATES B-CELL B-CELL INTERACTIONS. MAY BE INVOLVED IN
FT	DOMAIN 205 224	CC	THE LOCALIZATION OF B-CELLS IN LYMPHOID TISSUES. BINDS STYLYLATED
FT	DOMAIN 36 117	CC	GLYCOPROTEINS: ONE OF WHICH IS CD45.
FT	DOMAIN 123 175	CC	-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
FT	DISULFID 43 110	CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
FT	DISULFID 57 65	CC	HERE; ARE PRODUCED BY ALTERNATIVE SPLICING.
FT	CARBOHYD 90 90	CC	-1- TISSUE SPECIFICITY: B-LYMPHOCYTES.
FT	CARBOHYD 99 99	CC	-1- PTM: PHOSPHORYLATED BOTH ON THREONINE/SERINE AND TYROSINE.
FT	SEQUENCE 224 AA; 24830 MW; 60C88716D84600D2 CRC64;	CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
SQ		CC	-1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
		CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
		CC	-1- DATABASE: NAME=PROW; NOTE=CD guide CD22 entry;
	Query Match 8.0%; Score 100; DB 1; Length 224;	CC	WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd22.htm"
	Best Local Similarity 31.3%; Pred. No. 0.17;	CC	-----
	Matches 26; Conservative 15; Mismatches 38; Indels 4; Gaps 3;	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
QY	8 VEGELSGVTIKPLPEWH--VRVYLCEMAGSGTCGVVSTNFIKAEYGRVTKQYP 65	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
Db	31 VAGPVGSLVQCYRKEHRLTNKFWCRP-PQILRCDKIVETKG-SAGKRGVSRVSDSP 88	CC	the European Bioinformatics Institute. There are no restrictions on its
QY	66 RKNLFLVEVQLTESDGSVYACG 88	CC	use by non-profit institutions as long as its content is in no way
Db	89 ANLSFTVTLENLTEDAGTYWCG 111	CC	modified and this statement is not removed. Usage by and for commercial
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
		CC	or send an email to license@isb-sib.ch).
		CC	-----
RESULT 8		CC	EMBL; X59350; CAA42006.1; -.
CD22_HUMAN		DR	EMBL; X59350; CAA42006.1; -.
ID	CD22_HUMAN STANDARD; PRT: 847 AA.	DR	EMBL; X52785; CAA36988.1; -.
AC	P20273; Q01665; Q92872; O95699; O95701; O95702; O95703;	DR	EMBL; AB012997; BAA36566.1; -.
DT	01-FEB-1991 (Rel. 17, Created)	DR	EMBL; AB013003; BAA36572.1; -.
DT	15-JUL-1999 (Rel. 38, Last sequence update)	DR	EMBL; AB013004; BAA36573.1; -.
DT	01-MAR-2002 (Rel. 41, Last annotation update)	DR	EMBL; AB013006; BAA36575.1; -.
DE	B-cell receptor CD22 precursor (Leu-14) (B-lymphocyte cell adhesion	DR	PIR; A35648; A35648.
DE	molecule) (BL-CAM).	DR	PIR; JH0371; JH0371.
GN	CD22.	DR	MIM; 107266; -.
OS	Homo sapiens (Human).	DR	InterPro: IPR003006; Ig_MHC.
		DR	InterPro: IPR003598; Ig_c2.

DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 7.
DR SMART: SM00410; Ig_like; 2.
DR SMART: SM00408; IGG2; 4.
KW Glycoprotein; Cell adhesion; Transmembrane; Signal; B-cell;
KW Immunoglobulin domain; Alternative splicing; Phosphorylation;
KW Repeat; Polymorphism.

FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 847 B-CELL RECEPTOR CD22.
FT DOMAIN 20 687 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 688 706 POTENTIAL.
FT DOMAIN 707 847 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 138 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 154 226 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 258 316 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 346 403 IG-LIKE C2-TYPE DOMAIN 3.
FT DOMAIN 435 491 IG-LIKE C2-TYPE DOMAIN 4.
FT DOMAIN 522 578 IG-LIKE C2-TYPE DOMAIN 5.
FT DOMAIN 609 666 IG-LIKE C2-TYPE DOMAIN 6.
FT DISULFID 39 167 BY SIMILARITY.
FT DISULFID 44 102 BY SIMILARITY.
FT DISULFID 161 219 BY SIMILARITY.
FT DISULFID 265 309 BY SIMILARITY.
FT DISULFID 353 396 BY SIMILARITY.
FT DISULFID 442 484 BY SIMILARITY.
FT DISULFID 529 571 BY SIMILARITY.
FT DISULFID 616 659 BY SIMILARITY.
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 363 363 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 479 479 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 574 574 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 241 417 MISSING (IN ISOFORM CD22-ALPHA).
FT VARSPLIC 806 847 DYNV1PDPFDEGHIYSELIQGVGERQAOENVYVILK
H -> TMTSFQIFQMRGRITQS (IN ISOFORM
CD22-ALPHA).
FT VARIANT 152 152 Q -> E (OBSERVED WITH A marginally higher
FREQUENCY IN PATIENTS WITH SYSTEMIC LUPUS
ERYTHEMATOSUS).
FT VARIANT 664 664 /FTid=VAR_003913.
FT VARIANT 669 669 S -> G.
FT VARIANT 669 669 R -> C.
FT VARIANT 745 745 /FTid=VAR_003915.
FT CONFLICT 486 486 /FTid=VAR_003916.
FT CONFLICT 639 639 A -> R (IN REF. 1).
FT CONFLICT 788 789 Y -> H (IN REF. 1).
FT CONFLICT 788 789 PD -> RT (IN REF. 1).
SQ SEQUENCE 847 AA; 95347 MW; 38EB0BE4220ACA2D CRC64;

Query Match 7.7%; Score 96.5; DB 1; Length 847;
Best Local Similarity 20.7%; Pred. No. 1.7;
Matches 42; Conservative 31; Mismatches 69; Indels 61; Gaps 8;

Qy 1 RILPEVKVEGELGGSVTKCPLEPMHVRIYLCREMAAGSGTCGTWVSTNFIKAEYKGRVT 60
Db 247 KVTFSDAIVRE-GDSVTMTC-----EVSSSNPEYTTVSW-----LKDGTG 285
Qy 61 LKQYPRKNLFVETQLTESDSGVYACGAGNMTRDGTQKVTLMVHSEYBPSNWE----- 115
Db 286 LK---KQNTFTLNREVTKDQSGYKCCQVNDVGPGRSEEVFLQVYAPPESTVQILHSP 342
Qy 116 -----QPMPTPKWF-----HLPYLFQMPAYASSSKFVT-R 145
Db 343 AVEGSOVEFLCMLANPLPTNYTWYHNGKEMQGTKEKVHPIKI--LPWHAGTYSVCAEN 400

Qy 146 VITPAQRGKVPVPHSSPTTQIT 168
Db 401 ILGTGORGFGAELDVQYPPKVT 423
RESULT 9
ST50_YEAST STANDARD; PRT; 346 AA.
AC P25344;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE STE50 protein
GN STE50 OR YCL032W OR YCL32W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91377317; PubMed=1897318;
RA Rad M.R., Luetzenkirchen K., Xu G., Kleinhans U., Hollenberg C.P.;
RT "The complete sequence of a 11,953 bp fragment from CIG on chromosome
RT III encompasses four new open reading frames.";
RL Yeast 7:533-538(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93156679; PubMed=1494345;
RA Rad M.R., Xu G., Hollenberg C.P.;
RT "STE50, a novel gene required for activation of conjugation at an
RT early step in mating in Saccharomyces cerevisiae.";
RL Mol. Gen. Genet. 236:145-152(1992).
CC -!- FUNCTION: INVOLVED IN GROWTH ARREST DURING CONJUGATION. MAY
CC INTERACT WITH THE G PROTEIN ALPHA SUBUNIT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z11116; CAA77462.1; -.
DR EMBL: X59720; CAA42384.1; -.
DR PIR: S19360; S19360.
DR PIR: S17475; S17475.
DR SGD: S0000537; STE50.
DR InterPro: IPR000159; RA.
DR Pfam: PF00788; RA; 1.
DR SMART: SM00314; RA; 1.
KW Growth arrest; Pheromone response.
SQ SEQUENCE 346 AA; 38970 MW; EFC7D14F2940C779 CRC64;

Query Match 7.4%; Score 93; DB 1; Length 346;
Best Local Similarity 22.2%; Pred. No. 1.1;
Matches 56; Conservative 43; Mismatches 101; Indels 52; Gaps 10;

Qy 2 ILPEVKV-----EGELGGSVTKCPLEPMHVRIYLCREMAAGSGTCGTWV---STTNF 50
Db 69 LLEELCLQDCQDCLDGLDNLKAKFKILINKMRDSKLEWKDDKTDQEDNIIVLNLYTTSA 128
Qy 51 IKAEYKGRVTLKOYPRKNLFVETQLTESDSGVYACG-----AGMNTDRGKTQKVTLN 104
Db 129 KLQEFQ-----SQYTLRMDVLDVMKTSSTSSSPINTHGVSTTVPPSSNNIIPSSDGVSL 183
Qy 105 -----VHSEYBPSNWEQPMPTPKFHLPLFQMPAYASSSKFVTRVTTTAQRGKVP 157
Db 184 QTDYFTVTHNRQSPSRRESPT-----VFQPSLSHKSCL-----HKDSKNKVP 228
Qy 158 VHSSSTTQITTHPRYSRASSVAGDKPRFLPSTTAKISALEGLLKQPOTPSYNNHTRLH 217

Db 229 I-----STNQSHPSAVSTANT-PGSPNEALKQLRASKEDSCERILKNAMKRRHLADQDW 282
 QY 218 RORAL--DYGSO 227
 Db 283 RQVLVICYGQO 294

RESULT 10
 NCRL_HUMAN
 ID NCRL_HUMAN STANDARD; PRT; 2440 AA.
 AC 075376; O9UPV5; O9U018;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nuclear receptor co-repressor 1 (N-CoR1) (N-CoR).
 GN NCRL OR KIAA1047.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=98393736; PubMed=9724795;
 RA Wang J., Hoshino T., Redner R.L., Kajigaya S., Liu J.M.;
 RT "ETO, fusion partner in t(8;21) acute myeloid leukemia, represses
 transcription by interaction with the human N-CoR/MSIN3/HDAC1
 complex.";
 RT Proc. Natl. Acad. Sci. U.S.A. 95:10860-10865(1998).
 RL [2]
 RN SEQUENCE OF 782-2440 FROM N.A.
 RP TISSUE=Brain;
 RX MEDLINE=99397452; PubMed=10470851;
 RA Kikuno R., Nagase T., Ishikawa K.-I., Hirotsawa M., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:197-205(1999).
 RN [3]
 RP SEQUENCE OF 974-2440 FROM N.A.
 RX MEDLINE=99375328; PubMed=1044336;
 RA Nagaya T., Chen K.-S., Fujieda M.,
 RA Horwitz K.B., Lupski J.R., Seo H.;
 RT "Localization of the human nuclear receptor co-repressor (hn-CoR) gene
 RT between the CMT1A and the SMS critical regions of chromosome
 RT 17p11.2.";
 RL Genomics 59:339-341(1999).
 CC -1- FUNCTION: MEDIATES THE TRANSCRIPTIONAL REPRESSION ACTIVITY OF SOME
 CC NUCLEAR RECEPTORS BY PROMOTING CHROMATIN CONDENSATION, THUS
 CC PREVENTING ACCESS OF THE BASAL TRANSCRIPTION.
 CC -1- SUBUNIT: FORMS A LARGE COREPRESSOR COMPLEX THAT CONTAINS SIN3A/B
 CC AND HISTONE DEACETYLASES HDAC1 AND HDAC2. THIS COMPLEX ASSOCIATES
 CC WITH THE THYROID (TR) AND THE RETINOIC ACID RECEPTORS (RAR) IN THE
 CC ABSENCE OF LIGAND.
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- DOMAIN: THE N-TERMINAL REGION CONTAINS REPRESSION FUNCTIONS THAT
 CC ARE DIVIDED INTO THREE INDEPENDANT REPRESSION DOMAINS (RD1, RD2
 CC AND RD3). THE C-TERMINAL REGION CONTAINS THE NUCLEAR RECEPTOR-
 CC INTERACTING DOMAINS THAT ARE DIVIDED IN TWO SEPARATE INTERACTION
 CC DOMAINS (ID1 AND ID2).
 CC -1- DOMAIN: THE TWO INTERACTION DOMAINS (ID) CONTAIN A CONSERVED
 CC SEQUENCE REFERRED TO AS THE CORNR BOX. THIS MOTIF IS REQUIRED AND
 CC SUFFICIENT TO PERMIT BINDING TO UNLIGANDED TR AND RARS. SEQUENCES
 CC FLANKING THE CORNR BOX DETERMINE NUCLEAR HORMONE RECEPTOR
 CC SPECIFICITY.
 CC -1- SIMILARITY: CONTAINS 1 SANT-A DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 CORNR BOX.
 CC -1- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF044209; AAC33550.1; -;
 DR EMBL; AB028970; BAA82999.1; -;
 DR EMBL; AB019524; BAA75814.1; -;
 DR MIM; 600849; -;
 DR InterPro; IPR001005; Myb_DNA_bind.
 DR Pfam; PF00249; myb_DNA-binding; 2.
 DR SMART; SM00395; SANT; 2.
 DR PROSITE; PS0030; MYB_3; 1.
 KW Nuclear protein; Transcription regulation; DNA-binding; Repressor;
 KW Coiled coil.
 FT DOMAIN 174 216 COILED COIL (POTENTIAL).
 FT DOMAIN 254 312 INTERACTION WITH SIN3A/B.
 FT DOMAIN 299 328 COILED COIL (POTENTIAL).
 FT DNA_BIND 437 482 SANT-A (POTENTIAL).
 FT DNA_BIND 620 670 MYB.
 FT DOMAIN 501 557 COILED COIL (POTENTIAL).
 FT DOMAIN 607 617 PRO-RICH.
 FT DOMAIN 988 1816 INTERACTION WITH ETO.
 FT DOMAIN 2055 2059 CORNR BOX OF ID1.
 FT DOMAIN 2263 2267 CORNR BOX OF ID2.
 FT DOMAIN 58 64 POLY-GLN.
 FT DOMAIN 593 603 POLY-ALA.
 FT DOMAIN 1032 1035 POLY-PRO.
 FT DOMAIN 1707 1712 POLY-ALA.
 FT DOMAIN 1952 1963 POLY-SER.
 FT CONFLICT 1014 1014 L -> V (IN REF. 2).
 FT CONFLICT 1508 1509 PP -> SS (IN REF. 2).
 FT CONFLICT 1561 1561 W -> R (IN REF. 2).
 FT CONFLICT 1567 1567 Q -> H (IN REF. 2).
 SQ SEQUENCE 2440 AA; 270263 MW; 60A4D7964D00EDAB CRC64;
 Query Match 7.48; Score 93; DB 1; Length 2440;
 Best Local Similarity 23.08; Pred. No. 12;
 Matches 50; Conservative 24; Mismatches 79; Indels 64; Gaps 10;
 QY 14 GSVTIKCP-----LPENHVIYLCREMGSGTC---GTVVS---TTNFKAIEYKGR 58
 Db 912 GSILVSSPLKPNLDPQLQHRAAVDPWVSVCTPCNIPGTVPVSGYALYQRIKAMHESA 971
 QY 59 VTAKQYPRKNLFLVEVTQLTESDSGVYACGAGNMTRDGTOKVTLNVHSEYFSEWEE-OP 117
 Db 972 LLEEQRQHQIDECRSSTS-----PCGTSKSPNR-----EWEIOP 1009
 QY 118 MP-----ETPKWFHLPYLFQMPAYASSKFEVTRVTPAQRGVPPVHHSPFTTQITHRPR 172
 Db 1010 APhOLITNLPEGVRLP-----TTRPTRPP-----PPLIPSSKTTVASEKPS 1050
 QY 173 VSRASSVAGDKPRTFLPSTTASKISALEGLLKPKQTPS 209
 Db 1051 FIMGSGISQGTPTGYL--TSHNQASYQTETPKPSVGS 1085
 RESULT 11
 PLK_CHICK
 ID PLK_CHICK STANDARD; PRT; 355 AA.
 AC P07354;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Proteoglycan link protein precursor (Cartilage link protein) (LP).
 GN CRT11.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus
 OX NCBI_TaxID=9031;

CC MYOSIN LIGHT CHAIN, WHICH LEADS TO THE FORMATION CALMODULIN/
 CC MLCK SIGNAL TRANSDUCTION COMPLEXES WHICH ALLOW SELECTIVE
 CC TRANSDUCTION OF CALCIUM SIGNALS.
 CC -!- CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin
 CC light-chain] phosphate.
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST THREE ISOFORMS OF MLCK ARE PRODUCED
 CC BY ALTERNATIVE INITIATION, A NON-MUSCLE FORM (MLCK-210; WHICH IS
 CC THE LONGEST FORM); A SMOOTH-MUSCLE FORM (MLCK-108) AND TELOKIN (A
 CC C-TERMINAL SECT) WITH NO CATALYTIC ACTIVITY).
 CC -!- TISSUE SPECIFICITY: TELOKIN IS EXPRESSED IN GIZZARD, HEART, LUNG,
 CC INTESTINE, AND SKELETAL MUSCLE ALTHOUGH THE LEVELS OF THE
 CC EXPRESSION IN THE LATTER WERE MUCH LESS THAN THAT IN THE GIZZARD.
 CC -!- DOMAIN: TELOKIN BINDS CALMODULIN.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X52876; CAA37056.1; -;
 CC EMBL; X52876; CAA37057.1; -;
 CC EMBL; X52876; CAA37058.1; -;
 CC EMBL; M31048; AAA49069.1; -;
 CC EMBL; M14953; AAA69964.1; -;
 CC EMBL; M96655; AAA49083.1; -;
 CC EMBL; M88283; AAA48647.1; -;
 CC EMBL; M88284; AAB53768.1; -;
 CC PIR; A25810; A25810.
 CC PIR; A37099; A37099.
 CC PIR; S11652; S11652.
 CC HSSP; P56276; ITLK.
 CC InterPro; IPR000719; Euk_pkinase.
 CC InterPro; IPR003961; FN.III.
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR003598; Ig_C2.
 CC InterPro; IPR003600; Ig_Like.
 CC InterPro; IPR002290; Ser_thr_pkinase.
 CC Pfam; PF00041; fn3; 1.
 CC Pfam; PF00047; Ig; 9.
 CC Pfam; PF00069; pkinase; 1.
 CC SMART; SM00060; FN3; 1.
 CC SMART; SM00410; Ig_Like; 1.
 CC SMART; SM00408; IGC2; 8.
 CC SMART; SM00220; S_TKc; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC Transferrase; Serine/threonine-protein kinase; Calmodulin-binding;
 CC ATP-binding; Phosphorylation; Immunoglobulin domain; Repeat;
 CC Alternative initiation.
 CC CHAIN 1 1906
 FT MYOSIN LIGHT CHAIN KINASE, NON-MUSCLE
 FT ISOZYME (MLCK-210).
 FT MYOSIN LIGHT CHAIN KINASE, SMOOTH-MUSCLE
 FT ISOZYME (MLCK-108).
 FT TELOKIN.
 FT CHAIN 935 1906
 FT CHAIN 1750 1906
 FT INIT MET 935 935
 FT INIT MET 1750 1750
 FT DOMAIN 1330 1400
 FT FIBRONECTIN TYPE-III.
 FT DOMAIN 1453 1708
 FT PROTEIN KINASE.
 FT DOMAIN 1809 1876
 FT IG-LIKE C2-TYPE DOMAIN.
 FT NP_BIND 1459 1467
 FT ATP (BY SIMILARITY).
 FT BINDING 1482 1482
 FT ACT-SITE 1574 1574
 FT CALMODULIN AUTOINHIBITION (AM13) REGION
 FT DOMAIN 1716 1728
 FT (POTENTIAL).
 FT CALMODULIN RECOGNITION (RS20) REGION
 FT DOMAIN 1730 1749
 FT (POTENTIAL).

FT DOMAIN 1317 1364
 FT MOTIF IA.
 FT MOTIF IB.
 FT 4 X REPEATS, MOTIF IIA.
 FT REPEAT 660 1833
 FT IIA-1.
 FT REPEAT 660 676
 FT IIA-2.
 FT REPEAT 774 774
 FT IIA-3.
 FT REPEAT 1107 1123
 FT IIA-4.
 FT REPEAT 1817 1833
 FT 5 X REPEATS, MOTIF IIB.
 FT DOMAIN 693 1866
 FT REPEAT 693 708
 FT IIB-1.
 FT REPEAT 791 807
 FT IIB-2.
 FT REPEAT 1140 1156
 FT IIB-3.
 FT REPEAT 1281 1297
 FT IIB-4.
 FT REPEAT 1851 1866
 FT IIB-5.
 FT DOMAIN 970 1226
 FT 4 X REPEATS, MOTIF III.
 FT REPEAT 970 987
 FT IIB-1.
 FT REPEAT 999 1016
 FT IIB-2.
 FT REPEAT 1061 1078
 FT IIB-3.
 FT REPEAT 1209 1226
 FT IIB-4.
 FT CALMODULIN-BINDING.
 FT DOMAIN 1700 1763
 FT POLY-GLU.
 FT MOD_RES 1748 1748
 FT PHOSPHORYLATION.
 FT MOD_RES 1762 1762
 FT PHOSPHORYLATION.
 FT CONFLICT 1439 1439
 FT R -> Q (IN REF. 4).
 FT SEQUENCE 1906 AA; AD7D8A3B69EE3363 CRC64;
 SQ
 Query Match 7.38; Score 92; DB 1; Length 1906;
 Best Local Similarity 22.68; Pred. No. 11;
 Matches 49; Conservative 28; Mismatches 48; Indels 92; Gaps 12;
 QY 12 LGGSVTIKCP---LPMEHVRIYLCREMAGSGTCGTVVSTTNFIKAEYKGRVTLKQYPR-- 66
 Db 19 LSPVPAEAPFTLPPNIRVOL-----GA-----TARFEKGV--RGYPBPQ 58
 QY 67 -----KNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTLN 104
 Db 59 ITWYRNHGPLDEGHVVDHSIRGIFSLVKGQGDGSKYTCEA---ANDGVQVIVE 115
 QY 105 VHSEYEPSEWEEQMPETPKFHLPLFQMPAYASSKFFVTVTPPAQRKGVPPVHHSSPT 164
 Db 116 LTVE-----GNSLKKYSLP-----SSAK-----TPGRLSVPPVE----- 145
 QY 165 TQITHRPVRSASSVAGDKPRTLPSTASKISALEG 201
 Db 146 ----HRP-----SINGESPPEF--AIKPNRVVVRREG 170
 RESULT 13
 ID FPL_MYTCO STANDARD; PRT; 872 AA.
 AC Q25434;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Adhesive plaque matrix protein precursor (Foot protein 1) (MCFP1).
 GN FPL.
 OS Mytilus coruscus (Sea mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=42192;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Foot;
 RC MEDLINE=96394686; PubMed=8798340;
 RA Inoue K., Takeuchi Y., Takeyama S., Yamahara E., Yamazaki F., Odo S.,
 RA Harayama S.;
 RT "Adhesive protein cDNA sequence of the mussel Mytilus coruscus and
 RT its evolutionary implications.";
 RL J. Mol. Evol. 43:348-356(1996).
 CC -!- FUNCTION: PROVIDES ADHESIVENESS TO THE MUSSEL'S FOOT. MUSSELS
 CC PRODUCE ONE OF THE STRONGEST WATER INSOLUBLE GLUE. THE MUSSEL'S
 CC ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A
 CC FIBROUS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.

Tue Oct 29 09:47:51 2002

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04374; AAA43039.1; -.
DR FIR; J00102; RWPPEM.
DR MEROPS; C21.001; -.
DR InterPro; IPR001788; RNA_dep_RNApol2.
DR InterPro; IPR000606; Viral_helicase1.
DR Pfam; PF00978; RNA_dep_RNApol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
DR Transferase; RNA-directed RNA polymerase; Polyprotein; ATP-binding.
KW NP_BIND 965 972 ATP (By SIMILARITY).
FT SEQUENCE 1839 AA; 204731 MW; FD8DC1F5115E7861 CRC64;
SQ
```

Query Match 7.2%; Score 90.5; DB 1; Length 1839;
Best Local Similarity 26.0%; Pred. No. 14;
Matches 46; Conservative 16; Mismatches 80; Indels 35; Gaps 7;

```
Qy 58 RYTLKQYPRKNL-----FLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVH-SEYEP 111
Db 493 RIELQRLPLMSLIPKPKIVPLLSLLSSPTIYIHFFQAQTPQQLHDNYHLHLPSRFEL 552
Qy 112 SWEEQPMPTPKWFHLYPLFQMPAYASSKFEVTRVTPAQRCKYPPVHH----- 160
Db 553 SWTLQSYHVTQAQSFLLPLLPAPTQAQAS-----NPAPR---PPAFHAIPLPQPFSTS 602
Qy 161 SSFTTQ-----ITHRPRVSRASSVAGDK-PRTFPLPSTTASKISALEGLLKQPTPS 209
Db 603 SSFPLQEPFLSPHLIHPPLTREPSPLNGACDSALLPSTAAMTSAEHPTPLNPPTPS 659
```

Search completed: October 28, 2002, 17:31:27
Job time : 10.6613 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 3.33065 Seconds
(without alignments)
1371.778 Million cell updates/sec

Title: US-09-135-238B-2_COPY_273_390
Perfect score: 616
Sequence: 1 KRAVERKALSRARRLAVR.....HQPAMMEDSDDDYINPVA 118

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.5	16.0	116	1 STP2_RAT	P11101 rattus norv
2	87	14.1	2045	1 CBP1_CAEEL	P34545 caenorhabdi
3	85.5	13.9	2205	1 POLN_RUBVT	P13889 rubella vir
4	85.5	13.9	3866	1 HRX_MOUSE	P55200 mus musculu
5	81	13.1	549	1 FAST_HUMAN	Q14296 homo sapien
6	81	13.1	550	1 SRCB_HUMAN	Q14247 homo sapien
7	80	13.0	466	1 BLAR_MOUSE	P34971 mus musculu
8	80	13.0	466	1 BLAR_RAT	P18090 rattus norv
9	79.5	12.9	525	1 BTB2_HUMAN	Q9bx70 homo sapien
10	79	12.8	295	1 PLS3_HUMAN	Q9nr76 homo sapien
11	78.5	12.7	425	1 POXN_DROME	P23758 drosophila
12	78	12.7	863	1 MCX4_HUMAN	P33991 homo sapien
13	77.5	12.6	415	1 SYNI_CANFA	O62732 canis famli
14	77.5	12.6	942	1 M3KE_MOUSE	Q9wul6 mus musculu
15	77	12.5	248	1 ICT3_HSV11	P36313 herpes simp
16	77	12.5	577	1 CST2_HUMAN	P33240 homo sapien
17	77	12.5	938	1 EBNA_EBV	P03203 epstein-bar
18	76.5	12.4	258	1 UL49_HSVBC	P30022 bovine herp
19	76.5	12.4	444	1 FXF2_HUMAN	Q12947 homo sapien
20	76.5	12.4	587	1 FOLC_MOUSE	P48760 mus musculu
21	76.5	12.4	862	1 MCX4_MOUSE	P49717 mus musculu
22	76	12.3	1185	1 DRPL_HUMAN	P54259 homo sapien
23	75.5	12.3	988	1 CLC1_HUMAN	P35523 homo sapien
24	75.5	12.3	1616	1 SLAP_BACCI	P35824 bacillus cl
25	75	12.2	285	1 HXA4_MOUSE	P06798 mus musculu
26	75	12.2	335	1 PLS1_RAT	P58195 rattus norv
27	75	12.2	495	1 SMA6_MOUSE	Q35182 mus musculu
28	75	12.2	543	1 VP61_NPVAC	Q03209 autographa
29	75	12.2	994	1 CLC1_MOUSE	Q64347 mus musculu
30	75	12.2	1271	1 BCR_HUMAN	P11274 homo sapien
31	74.5	12.1	294	1 YQ19_CAEEL	Q09507 caenorhabdi
32	74.5	12.1	308	1 CC40_CAEEL	P34804 caenorhabdi
33	74.5	12.1	366	1 IHA_RAT	P17490 rattus norv

RESULT 1

ID	STP2_RAT	STANDARD:	PRT:	116 AA.
AC	P11101:			
DC	01-JUL-1989 (Rel. 11, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DE	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	Nuclear transition protein 2 (Tp-2).			
GN	TNP2.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89263797; PubMed=2726489;			
RA	Luerssen H., Maier W.M., Hoyer-Fender S., Engel W.;			
RT	"The nucleotide sequence of rat transition protein 2 (TP2) cDNA."			
RL	Nucleic Acids Res. 17:3585-3585(1989).			
RN	[2]			
RP	SEQUENCE OF 91-116.			
RX	MEDLINE=87326384; PubMed=3307778;			
RA	Cole K.D., Kistler W.S.;			
RT	"Nuclear transition protein 2 (TP2) of mammalian spermatids has a very basic carboxyl terminal domain."			
RL	Biochem. Biophys. Res. Commun. 147:437-442(1987).			
RN	[3]			
RP	ZINC-BINDING			
RX	MEDLINE=92028893; PubMed=1930189;			
RA	Baskaran R., Rao M.R.S.;			
RT	"Mammalian spermatid specific protein, TP2, is a zinc metalloprotein with two finger motifs."			
RL	Biochem. Biophys. Res. Commun. 179:1491-1499(1991).			
CC	!- FUNCTION: IN THE ELONGATING SPERMATIDS OF MAMMALS, THE CONVERSION OF NUCLEOSOMAL CHROMATIN TO THE COMPACT, NONNUCLEOSOMAL FORM FOUND IN THE SPERM NUCLEUS IS ASSOCIATED WITH THE APPEARANCE OF A SMALL SET OF BASIC CHROMOSOMAL TRANSITION PROTEINS.			
CC	!- SUBCELLULAR LOCATION: Nuclear.			
CC	!- SIMILARITY: STRONG, TO OTHER MAMMALIAN SPERMATID NUCLEAR TRANSITION PROTEINS 2.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
DR	EMBL; X14776; CAA32882.1; .			
DR	PIR; A26834; A26834.			
DR	PIR; S04094; S04094.			
DR	InterPro; IPR000678; TP2.			
DR	Pfam; PF01254; TP2; 1.			
DR	PROSITE; PS00970; TP2_1; 1.			
DR	PROSITE; PS00971; TP2_2; 1.			


```

DR Pfam; PF00978; RNA_dep_RNApol2; 1.
DR SMART; SM00506; Alpp; 1.
KW Polyprotein; Nonstructural protein.
FT CHAIN 1 ? NONSTRUCTURAL PROTEIN NSP1.
FT CHAIN 2 ? NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 3 ? NONSTRUCTURAL PROTEIN NSP3.
FT CHAIN 4 ? NONSTRUCTURAL PROTEIN NSP4.
SQ SEQUENCE 2205 AA; 240220 MW; 15A95F55E34C0B3 CRC64;

Query Match 13.9%; Score 85.5; DB 1; Length 2205;
Best Local Similarity 31.4%; Pred. No. 14;
Matches 38; Conservative 9; Mismatches 43; Indels 31; Gaps 7;

QY 9 ALSRRRLAVRM--RALESSQPRGSPRRSONNIYSACPRRA-----RGADAAGT---G 59
Db AFDAVAVTAVRAGRQSAASAPPGDPPP-----PRRRSRSHSDARGTTPPA 748

QY 60 EAPVPGCAPLPPLQVSE--SPWLHAPS-----LKTSCYVSLYHQPAAMMEDSDS 110
Db FARDDPPPPAPPPAPPRAGDVPVPIAGPADRADAELEVACEPSG---PPTSTRADPDS 805

QY 111 D 111
Db 806 D 806

RESULT 4
HRX_MOUSE STANDARD; PRT; 3866 AA.
AC P55200;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Zinc finger protein HRX (ALL-1) (Fragment).
GN MLL OR HRX OR ALL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=C57BL/6J, AND B6/CBA; TISSUE=Spleen, and Lung;
RX MEDLINE=93317679; PubMed=8327517;
RA Ma Q., Alder H., Nelson K.K., Chatterjee D., Gu Y., Nakamura T.,
RA Canani E., Croce C.M., Siracusa L.D., Buchberg A.M.;
RT "Analysis of the murine All-1 gene reveals conserved domains with
RT human ALL-1 and identifies a motif shared with DNA
RT methyltransferases.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6350-6354(1993).
CC -!- FUNCTION: POSSIBLY ACTS AS A TRANSCRIPTIONAL REGULATORY FACTOR.
CC MAY REGULATE GENES INVOLVED IN SKELETAL FORMATION DURING
CC EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE TRANSCRIPTION FACTOR TRITHORAX FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BROMODOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SET DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 PHD-TYPE ZINC FINGERS.
CC -!- SIMILARITY: CONTAINS 1 CXXC-TYPE ZINC FINGER.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L17069; AAA62593.1; -.
CC MGD; MGI:96995; Mll.
CC InterPro; IPR001487; Bromodomain.
CC InterPro; IPR003889; FYrich_C.

```

```

DR InterPro; IPR003888; FYrich_N.
DR InterPro; IPR001965; PHD.
DR InterPro; IPR003636; PostSET.
DR InterPro; IPR000093; Recr.
DR InterPro; IPR001214; SET.
DR InterPro; IPR002857; Znf-CXXC.
DR Pfam; PF00628; PHD; 3.
DR Pfam; PF00856; SET; 1.
DR Pfam; PF02008; zf-CXXC; 1.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00542; FYRC; 1.
DR SMART; SM00541; FYRN; 1.
DR SMART; SM00249; PHD; 4.
DR SMART; SM00508; PostSET; 1.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS50014; BROMODOMAIN_2; 1.
DR PROSITE; PS50280; SET; 1.
KW DNA-binding; Nuclear protein; Zinc-finger; Metal-binding;
KW Transcription regulation; Alternative splicing; Polymorphism.
FT NON_TER 1 1
FT DNA_BIND 67 78 A.T HOOK (BY SIMILARITY).
FT DNA_BIND 115 125 A.T HOOK (BY SIMILARITY).
FT DNA_BIND 199 207 A.T HOOK (BY SIMILARITY).
FT ZN_FING 1044 1091 CXXC-TYPE.
FT ZN_FING 1330 1381 PHD-TYPE 1.
FT ZN_FING 1383 1432 PHD-TYPE 2.
FT ZN_FING 1465 1529 PHD-TYPE 3.
FT DOMAIN 1605 1650 BROMODOMAIN (DIVERGENT).
FT DOMAIN 3737 3846 SET.
FT DOMAIN 35 41 POLY-GLY.
FT DOMAIN 459 469 PRO-RICH.
FT DOMAIN 1231 1238 POLY-PRO.
FT DOMAIN 3533 3536 POLY-GLU.
FT DOMAIN 3693 3697 POLY-GLU.
FT VARSPIC 1503 1505 MISSING (IN ISOFORM 2).
FT VARIANT 1497 1497 K -> T.
SQ SEQUENCE 3866 AA; 420976 MW; ADFC55E14E806FID CRC64;

Query Match 13.9%; Score 85.5; DB 1; Length 3866;
Best Local Similarity 28.2%; Pred. No. 24;
Matches 31; Conservative 21; Mismatches 37; Indels 21; Gaps 6;

QY 1 KLAVERRRKALSRRLAVRMRA-LESSOR-----PRGSPRRSONNIYSACPRRARGADA 55
Db 1111 KKEKESKTKTEKESKESVSKPLPAQAKAPPPREAPAKKSS---SEPPRKPVEEK 1167

QY 56 AGTGEAPVPGCAPLPPLQVSESPWLHAPSLKTSCEYVSLYHQPAAM 105
Db 1168 SEEGGAPAPAPA---PEPKQVS-----APASRKSKQVS---QPAAVV 1204

RESULT 5
FAST_HUMAN STANDARD; PRT; 549 AA.
AC Q14296;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fas-activated serine/threonine kinase (EC 2.7.1.-) (FAST kinase).
GN FASTK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood lymphocytes;
RX MEDLINE=95378805; PubMed=7544399;
RA Tian O., Taupin J.-L., Elledge S., Robertson M., Anderson P.;
RT "Fas-activated serine/threonine kinase (FAST) phosphorylates TIA-1
RT during Fas-mediated apoptosis.";
RL J. Exp. Med. 182:865-874(1995).
CC -!- FUNCTION: FAS-MEDIATED APOPTOSIS IS CHARACTERIZED BY FASTK-

```


QY 78 SESPWLHAPSLKTCSEYV-----SLYHQPAAMME 106
 DB 62 SAAPFLPLGVPVSGLEFLVQIDQILHOKAERVE 95

RESULT 11
 POXN_DROME
 AC P23758;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 GN POX-N.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=92208941; PubMed=1348214;
 RA Dambly-Chaudiere C., Jamet E., Burri M., Bopp D., Basler K.,
 RA Hafen E., Dumont N., Spielmann P., Ghysen A., Noll M.;
 RT "The paired box gene pox-neuro: a determinant of poly-innervated
 sense organs in Drosophila.";
 RL Cell 69:159-172(1992).
 RN [2]
 RN SEQUENCE OF 1-35 FROM N.A.
 RX MEDLINE=90059940; PubMed=2573516;
 RA Bopp D., Jamet E., Baumgartner S., Burri M., Noll M.;
 RT "Isolation of two tissue-specific Drosophila paired box genes, Pox
 meso and Pox-neuro.";
 RL EMBO J. 8:3447-3457(1989).
 RN [3]
 RN FUNCTION.
 RX MEDLINE=94121914; PubMed=8292359;
 RA Nottebohm E., Usui A., Therianos S., Kimura K., Dambly-Chaudiere C.,
 RA Ghysen A.;
 RT "The gene poxn controls different steps of the formation of
 chemosensory organs in Drosophila.";
 RL Neuron 12:25-34(1994).
 RN [4]
 RN FUNCTION.
 RX MEDLINE=97326936; PubMed=9183748;
 RA Awasaki T., Kimura K.;
 RT "pox-neuro is required for development of chemosensory bristles in
 Drosophila.";
 RL J. Neurobiol. 32:707-721(1997).
 CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT SPECIFIES POLY-INNERVATED
 CC ORGANS (CHEMOSENSORY BRISTLE). ALSO CONTROLS THE NUMBER OF
 CC NEURONS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: CENTRAL AND PERIPHERAL NERVOUS SYSTEM.
 CC -!- SIMILARITY: CONTAINS A PAIRED BOX DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M86927; AAA28832.1; -;
 CC EMBL; X58917; CAA41721.1; -;
 CC PIR; S06951; S06951.
 CC HSP; P26367; 6PAX.
 CC FlyBase; FBgn0003130; Pox-n.
 CC InterPro; IPR001523; Paired_box.
 CC Pfam; PF00292; PAX; 1.
 CC PRINTS; PR00027; PAIREDBOX.

DR SMART: SM00351; PAX; 1.
 DR PROSITE: PS00034; PAIRED_BOX; 1.
 KW Paired box; Developmental protein; Nuclear protein;
 KW Transcription regulation; Neurogenesis; Differentiation.
 FT DOMAIN 5 130 PAIRED_BOX.
 FT DOMAIN 146 157 POLY-ALA.
 FT DOMAIN 173 176 POLY-PRO.
 FT DOMAIN 281 287 POLY-ALA.
 SQ SEQUENCE 425 AA; 44387 MW; EC76C6494A06DC15 CRC64;
 Query Match 12.7%; Score 78.5; DB 1; Length 425;
 Best Local Similarity 32.4%; Pred. No. 9.7; Indels 2;
 Matches 23; Conservative 9; Mismatches 30; Gaps 2;
 QY 38 SNNIYSACPRRARGAAGTGEAPVPGFAPLPAPLQVSESPWLHAPSLKTCSEYVSL 97
 DB 142 SQNAAAAAHHAGSG--PSNGYGGQAPPVTV-----APPTPAATPSIAR 192
 QY 98 YHQPAAMMEDS 108
 DB 193 YAKPPALMMNS 203

RESULT 12
 MCM4_HUMAN
 ID MCM4_HUMAN STANDARD; PRT; 863 AA.
 AC P33991; Q99658;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA replication licensing factor MCM4 (CDC21 homolog) (P1-CDC21).
 GN MCM4 OR CDC21
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Hu B.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE OF 1-712 FROM N.A.
 RX MEDLINE=98126438; PubMed=9465298;
 RA Connelly M.A., Zhang H., Kieleczawa J., Anderson C.W.;
 RT "The promoters for human DNA-PKcs (PRKDC) and MCM4: divergently
 RT transcribed genes located at chromosome 8 band q11.";
 RL Genomics 47:71-83(1998).
 RN [3]
 RN SEQUENCE OF 1-23 FROM N.A.
 RX MEDLINE=97430835; PubMed=9284934;
 RA Ladenburger E.M., Fackelmayr F.O., Hamelster H., Knippers R.;
 RT "MCM4 and PRKDC, human genes encoding proteins MCM4 and DNA-PKcs, are
 RT close neighbours located on chromosome 8q12->q13.";
 RL Cytogenet. Cell Genet. 77:268-270(1997).
 RN [4]
 RN SEQUENCE OF 440-863 FROM N.A.
 RX TISSUE=Cervix;
 CC MEDLINE=94089373; PubMed=8265339;
 CC Hu B., Burkhardt R., Schulte D., Musahl C., Knippers R.;
 CC "The P1 family: a new class of nuclear mammalian proteins related to
 CC the yeast Mcm replication proteins.";
 CC Nucleic Acids Res. 21:5289-5293(1993).
 CC -!- FUNCTION: INVOLVED IN THE CONTROL OF DNA REPLICATION.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).


```

DR   EMBL; X14112; -: NOT_ANNOTATED_CDS.
DR   PIR; JQ1682; JQ1682.
DR   KW Repeat.
FT   DOMAIN          161      175      5 X 3 AA TANDEM REPEATS OF A-T-P.
SQ   SEQUENCE        248 AA;  26185 MW;  AD334AF0C9CABA65 CRC64;
      Query Match      12.5%;  Score 77;  DB 1;  Length 248;
      Best Local Similarity 25.0%;  Pred. No. 7.5;
      Matches 25;  Conservative 11;  Mismatches 36;  Indels 28;  Gaps 4;

QY   28  QPRGSRPRP-----SQNNIYSACPRRARGAAGTGEAPYGGCAPLPAPLQV 77
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   5  RRHRGPRRRPGTCAVPTAQSQVTSI-PNSEPAVRSAPAAAPPPAGGPPPSCLLL 63

QY   78  SESPLWHAESLTKTCYVSYSLYHQPAAMMEDSDSDYINVP 117
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   64  RO-WLHVPE-----SASDDDDDDDDWDPDSP 86

```


Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	98.5	16.0	116	2	S04094	transition protein	
2	98.5	14.4	526	2	D75391	Algp-related prote	
3	88.5	14.1	2027	2	S61023	hypothetical prote	
4	87	14.1	2056	2	G88564	protein R10E11.1	
5	86.5	14.0	113	2	S57667	transition protein	
6	85.5	13.9	2205	1	MWVRN	nonstructural poly	
7	85.5	13.9	3869	2	A48205	All-1 protein +GRE	
8	82.5	13.4	2115	2	S38480	nonstructural prot	
9	81.5	13.2	528	2	B75310	conserved hypothet	
10	81	13.1	549	2	I37386	fas-activated seri	
11	81	13.1	550	2	A48063	mammary tumor/squa	
12	81	13.1	553	2	C75318	hypothetical prote	
13	80.5	13.1	296	2	T12469	hypothetical prote	
14	80.5	13.1	381	2	T45482	hypothetical prote	
15	80.5	13.1	577	2	T09024	proline-rich prote	
16	80	13.0	464	2	S25591	beta-1-adrenergic	
17	80	13.0	466	2	S36794	beta-1-adrenergic	
18	79	12.8	285	1	A43556	homeotic protein H	
19	78.5	12.7	342	2	T23060	hypothetical prote	
20	78.5	12.7	425	2	A38153	paired box protein	
21	78.5	12.7	519	2	S14529	transition protein	
22	78	12.7	487	2	S61243	deoxyribonuclease	
23	78	12.7	863	1	S65954	replication licens	
24	78	12.7	1892	2	T18314	hypothetical prote	
25	78	12.7	2715	2	T13049	eyelid - fruit fly	
26	77	12.5	104	2	H96000	hypothetical prote	
27	77	12.5	248	1	JQ1682	infected cell prot	
28	77	12.5	314	2	B75588	Frah-related prote	
29	77	12.5	577	2	A40220	cleavage stimulati	


```

Db 1171 SEEGAPAPAPA-----PEPKQVS-----APASRKSKQVS---QPAVV 1207

RESULT 8
S38480
nonstructural protein - rubella virus
C:Species: rubella virus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 2
C:Accession: S38480
R:Gillam, S.
submitted to the EMBL Data Library, March 1993
A:Description: Nucleotide sequence of the nonstructural protein g
A:Reference number: S38480
A:Accession: S38480
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-2115 <GIL>
A:Cross-references: EMBL:X72393; NID:g410507; PIDN:CAA51087.1; PTF
C:Superfamily: rubella virus nonstructural polyprotein

Query Match 13.4%; Score 82.5; DB 2; Length 2115
Best Local Similarity 30.4%; Pred. No. 41;
Matches 38; Conservative 14; Mismatches 48; Indels 2

QY 2 RAVERRKALSRARRLAVRM--RALESSQPRGSPRPRSQNNIYSACPRA-----PRARRSQ
Db 693 RDLPRGTAFTANAVTAARAGPQGLAATSPPGDPPP-----PRARRSQ
QY 56 AGT-GEAPVGP--GAPLPAPLQVSE--SPWLHAPSLKTSCEYVSLYHQPA---
Db 743 RGTPTTAPVRDPPRPQPSPAPPRVGPVPTTTAEPADRARHAELEVVEYSPGPP
QY 107 DSDSD 111
Db 803 DPDS 807

RESULT 9
B75310
conserved hypothetical protein - Deinococcus radiodurans (strain 1
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 1
C:Accession: B75310
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterbach
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococo
A:Reference number: A75250; MUID:20036896
A:Accession: B75310
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-528 <WHI>
A:Cross-references: GB:AE002048; GB:AE000513; NID:g6459929; PIDN:
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2133
A:Map position: 1

Query Match 13.2%; Score 81.5; DB 2; Length 528;
Best Local Similarity 31.6%; Pred. No. 13;
Matches 30; Conservative 9; Mismatches 47; Indels 9

QY 2 RAVERRKALSRARRLAVRMRALESSQPRGSPRPRSQNNIYSACPRRARGADAAC
Db 437 RTGPRRARASRKPITAPAAANSAAAPPNSPTRKTWNSTPGWCPRSAATPSS
QY 62 PVP--GPGAPLPAPLQVSESPW--LHAPSLKTS 92
Db 497 PPRVGPFG-PEPTA----RRRGWGRPHRAPRWKPN 526

RESULT 10

```

I37386

fas-activated serine/threonine kinase - human

C:Species: Homo sapiens (man)

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000

C:Accession: I37386

R:Rian, O.; Taupin, J.; Elledge, S.; Robertson, M.; Anderson, P.

J. Exp. Med. 182, 865-874, 1995

A:Title: Fas-activated serine/threonine kinase (FAST) phosphorylates TIA-1 during Fas-me

A:Reference number: I37386; MUID:95378805

A:Accession: I37386

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-549 <RES>

A:Cross-references: EMBL:X86779; NID:g1006658; PIDN:CAA60448.1; PID:g1006659

C:Genetics:

A:Gene: fast

C:Superfamily: human fas-activated serine/threonine kinase

Query Match 13.1%; Score 81; DB 2; Length 549;
 Best Local Similarity 29.1%; Pred. No. 15;
 Matches 23; Conservative 5; Mismatches 23; Indels 28; Gaps 2;

QY 28 ORRGSPRRSQNNIYSACPRARGADAGTGEAPVPGCAPL-----70

DB 2 RRRGPGPRA-----PRTEGATCAGGGSWSPSNMLRVLLSAQTSAPRLSGLL 53

QY 71 ---PPAPLQVSESPWLHAP 86

DB 54 LIPPVQPCCLGPGSKWGRDP 72

RESULT 11

A48063

mammary tumor/squamous cell carcinoma-associated protein EMS1 - human

C:Species: Homo sapiens (man)

C:Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 29-Aug-1997

C:Accession: A48063

R:Schuring, E.; Verhoeven, E.; Litvinov, S.; Michalides, R.J.

Mol. Cell. Biol. 13, 2891-2898, 1993

A:Title: The product of the EMS1 gene, amplified and overexpressed in human carcinomas,

A:Reference number: A48063; MUID:93233650

A:Accession: A48063

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-550 <SCH>

A:Note: sequence extracted from NCBI backbone (NCBIP:129872)

C:Genetics:

A:Gene: GDB:EMS1

A:Cross-references: GDB:202962; OMIM:164765

A:Map position: 11q13-11q13

C:Superfamily: SH3 homology

F:499-546/Domain: SH3 homology <SH3>

Query Match 13.1%; Score 81; DB 2; Length 550;
 Best Local Similarity 27.0%; Pred. No. 15;
 Matches 37; Conservative 16; Mismatches 56; Indels 28; Gaps 6;

QY 1 KRAVERKALSRARRLAVMRRALESSOR-----PRGSPRRSQNNIYSACPR 48

DB 361 KEODERRAEARQMAKERQEQEARRKLEQARAKTQTPVSPAPQTEERLPSSPV 420

QY 49 RANGADAG-TGEAPVPGCAPLPPAPLQVSESPWLHAPSLKTSCEVY--SLY----- 98

DB 421 Y----EADASFKAELSLRGVSGTPEPEYISMEAADYREASSQGLAYATEAVYESAEPG 477

QY 99 HQPAAAMDESDSDYYIN 115

DB 478 HYP-----EDSYDEYEN 491

RESULT 12

C75318

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: C75318

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.

, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: C75318

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-553 <WHI>

A:Cross-references: GB:AE002044; GB:AE000513; NID:g6459872; PIDN:AAF11639.1; PID:g645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2090

A:Map position: 1

Query Match 13.1%; Score 81; DB 2; Length 553;
 Best Local Similarity 25.0%; Pred. No. 15;
 Matches 27; Conservative 16; Mismatches 37; Indels 28; Gaps 3;

QY 3 AVERKALSR---ARRLAVMRRALESSORPRSPRRSQNNIYSACPRRARGA----- 53

DB 176 AAPRREAPARQVQERSTATQTOIAAPKAPAPQPDSPSL----PRTLOALASDL 231

QY 54 -----DAAGTGEAPVPGCAPLPPAPLQVSESPWLHAP 86

DB 232 PDLPVLELLRLEWEQAAEQEPPAPAPRVPVAAALSPAPSPAAP 279

RESULT 13

T12469

hypothetical protein DKFzp564C1940.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999

C:Accession: T12469

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, June 1999

A:Reference number: 217525

A:Accession: T12469

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-296 <POU>

A:Cross-references: EMBL:AL080164

A:Experimental source: fetal brain; clone DKFzp564C1940

C:Genetics:

A:Note: DKFzp564C1940.1

Query Match 13.1%; Score 80.5; DB 2; Length 296;
 Best Local Similarity 28.3%; Pred. No. 9;
 Matches 43; Conservative 10; Mismatches 58; Indels 41; Gaps 6;

QY 6 RRKALSRARRLAVMR---ALESSORPRSPRRSQNNIYSACPRRARGA 53

DB 128 RRQRGRRLMRRLVHRLRWGLLPRTNTPARASEARSOVTPSAAPLEALDGGTGPAREGGA 187

QY 54 DAAGTGEAPVCP-GAPLP-----PAPLQVSESPWLHAPSLKTSCEYVSLYHQP----- 101

DB 188 VGSQDGEQAPPLTKAPLSASTSPAPTVPEAPG-PLPSLPLEPSLLSGVQVALLGRLL 246

QY 102 -----AAMDESDSDYYINVP 117

DB 247 PSLGPGPPTRSPPGPHTAVLALEDDVLLVP 278

RESULT 14

T46482

hypothetical protein DKFzp434A045.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46482

R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23036
A:Accession: T46482
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-381 <AAA>
A:Cross-references: EMBL:AL137629
A:Experimental source: adult testis; clone DKFZp434A045
C:Genetics:
A:Note: DKFZp434A045.1

Query Match 13.1%; Score 80.5; DB 2; Length 381;
Best Local Similarity 31.4%; Pred. No. 11;
Matches 32; Conservative 12; Mismatches 27; Indels 31; Gaps 8;
Qy 12 RRARLAVRRALESQRPRPRSONNIYSACPRRARGADAGTGEAPVPGP--GAP 69
| : | | | : : : | | | | | : | | | | : | | | | : |
Db 227 RKERSTAV-MRS-QPARLPQASPRP-----YSSVP-----AGS-EKPPKSSYNPP 269
| : | | | : : : | | | | | : | | | | : | | | | : |
Qy 70 LPPAPLQVSESPWLHAPSLKTCSEYVSLYHQPAAMMEDSDSD 111
| | | : | : | : | | | | | | | | | | : | | |
Db 270 LPPLKISTNSG---SPGPE-----YHQPCDKFEASKND 299

RESULT 15

T09024
proline-rich protein T27E11.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
C:Accession: T09024
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft, D.; et al.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16533
A:Accession: T09024
A:Molecule type: DNA
A:Residues: 1-577 <BEV>
A:Cross-references: EMBL:AL078579; GSPDB:GN00062; ATSP:T27E11.90
A:Experimental source: cultivar Columbia; BAC clone T27E11
C:Genetics:
A:Gene: ATSP:T27E11.90
A:Map position: 4
A:Introns: 26/1; 117/1; 338/1; 411/3; 430/2; 498/2

Query Match 13.1%; Score 80.5; DB 2; Length 577;
Best Local Similarity 28.8%; Pred. No. 17;
Matches 23; Conservative 10; Mismatches 32; Indels 15; Gaps 3;

Qy 30 PRGSPRPRSONNIYSACPRRARGADAGTGEAPVPGGAPLP-----PAPLQVSE 79
| | | | : | : | : | | | | | | | | | | : | | | : | |
Db 234 PSSGTPGDPSPLESPGPFSPSPPTGPDSPLESPGDPSPLESPGDPPLPSPGPHLYEK 293
| | | | : | : | : | | | | | | | | | | : | | | : | |
Qy 80 SPWL-HAPSLKTCSEYVSLY 98
| | | | : | : | : | | | | | | | | | | : | | | : | |
Db 294 NRWLIHFPSIK-----YLSVF 309

Search completed: October 28, 2002, 17:34:06
Job time : 10.1855 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:40 ; Search time 40.8871 Seconds
(without alignments)
1059.473 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLPLFLPVGALRL.....HQPAMMEDSDSDVINVPA 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	99.6	390	20	AA1980
2	2047	99.6	390	20	AA1981
3	2047	99.6	390	20	AA1982
4	1157	56.3	422	22	AAE05349
5	550	26.8	107	20	AA1980
6	410	20.0	97	21	AA1980
7	198	9.6	771	17	AA1980
8	195	9.5	771	22	AA1980
9	185	9.0	769	17	AA1980
10	185	9.0	769	22	AA1980
11	183.5	8.9	532	19	AA1980

12	182	8.9	607	20	AA1980
13	182	8.9	607	20	AA1981
14	182	8.9	746	17	AA1982
15	182	8.9	757	17	AA1983
16	182	8.9	757	22	AA1984
17	182	8.9	764	22	AA1985
18	182	8.9	764	22	AA1986
19	179	8.7	733	22	AA1987
20	166.5	8.1	584	12	AA1988
21	166.5	8.1	773	17	AA1989
22	166.5	8.1	773	22	AA1990
23	161	7.8	307	19	AA1991
24	156	7.6	584	12	AA1992
25	143.5	7.0	205	20	AA1993
26	142.5	6.9	211	21	AA1994
27	131	6.4	746	22	AA1995
28	129.5	6.3	332	20	AA1996
29	129.5	6.3	332	21	AA1997
30	129.5	6.3	332	21	AA1998
31	129.5	6.3	332	22	AA1999
32	129.5	6.3	332	22	AA2000
33	127.5	6.2	203	22	AA2001
34	127.5	6.2	313	22	AA2002
35	127.5	6.2	332	21	AA2003
36	122	5.9	298	20	AA2004
37	121	5.9	334	19	AA2005
38	120.5	5.9	301	20	AA2006
39	118.5	5.8	528	15	AA2007
40	117.5	5.7	476	18	AA2008
41	117.5	5.7	476	19	AA2009
42	117	5.7	453	14	AA2010
43	117	5.7	453	21	AA2011
44	117	5.7	1106	22	AA2012
45	117	5.7	1106	22	AA2013

ALIGNMENTS

RESULT 1
AA1980
ID AA1980 standard; Protein; 390 AA.
AC AA1980
DT 17-DEC-1999 (first entry)
XX Human Toso protein sequence.
DE Human Toso protein sequence.
KW Human Toso protein; target; drug screening; diagnosis; apoptosis;
KW apoptosis related disease.
XX Homo sapiens.
XX WO9950671-A2.
XX 07-OCT-1999.
XX 30-MAR-1999; 99WO-US06945.
XX 30-MAR-1998; 98US-0050861.
XX (RIGE-) RIGEL PHARM INC.
XX Payan D;
XX WPI; 1999-591379/50.
XX N-PSDB; AAZ25422.
XX Screening agents useful for modulating apoptosis and controlling
XX apoptosis related diseases
XX Claim 1; Fig 2a; 75pp; English.

XX The present invention describes a method of Screening for a bioactive
CC agent capable of binding a Toso protein. Also described a methods for:
CC (1) screening a bioactive agent capable of modulating activity of a
CC Toso cell-surface receptor, comprising adding a candidate bioactive
CC agent to a cell comprising a recombinant Toso nucleic acid, and
CC exposing the cells to an apoptotic agent that will induce apoptosis;
CC (2) modulating apoptosis comprising administering an exogenous
CC compound that binds Toso, to a cell; (3) identifying a cell containing
CC a mutant Toso gene, comprising determining it's sequence; (4)
CC identifying the Toso genotype, comprising determining the sequence of
CC at least one Toso gene; and (5) diagnosing an apoptosis related
CC condition, comprising measuring activity of Toso in a tissue, and
CC comparing to the activity from non-affected individual's tissue, where
CC a reduced activity of the patient indicates risk of an apoptosis related
CC condition. The methods are useful for identifying agents capable of
CC diagnosing and treating apoptosis related disease, their use for
CC modulating apoptosis, and methods for diagnosing the disease state.
CC The present sequence represents the human Toso protein for use in
CC methods from the invention.

XX SQ Sequence 390 AA;
Query Match 99.6%; Score 2047; DB 20; Length 390;
Best Local Similarity 99.7%; Pred. No. 1.8e-166;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDRWLWPLYFLPVSGALRLPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCT 60
|||
Db 1 MDFWLWPLYFLPVSGALRLPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCT 60
Qy 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTL 120
|||
Db 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTL 120
Qy 121 NVHSEYEPSWEQPMPTPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVHHSSP 180
|||
Db 121 NVHSEYEPSWEQPMPTPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVHHSSP 180
Qy 181 TQITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQTPSYNHHTRLHRQALD 240
|||
Db 181 TQITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQTPSYNHHTRLHRQALD 240
Qy 241 YGSQSGREGQGHILPTILGLFLALLGLVVKRAVERRKALSRARRLAVRRALESSQ 300
|||
Db 241 YGSQSGREGQGHILPTILGLFLALLGLVVKRAVERRKALSRARRLAVRRALESSQ 300
Qy 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360
|||
Db 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360
Qy 361 KTCEYVSLYHQPAAMMEDSDSDYINVPA 390
|||
Db 361 KTCEYVSLYHQPAAMMEDSDSDYINVPA 390

RESULT 2
AAV17496
ID AAV17496 standard; Protein; 390 AA.
XX AAV17496;
AC AAV17496;
XX 03-AUG-1999 (first entry)
XX Human Toso protein.
XX Toso protein; tumour necrosis factor mediated apoptosis inhibition;
KW TNF mediated apoptosis; T cell overactivity; autoimmune disease;
KW Sjogrens connective tissue disorder; transplant rejection; cancer.
XX Homo sapiens.
OS Homo sapiens.
XX FH Key Location/Qualifiers

FT Peptide 1..17
FT /label= signal
FT Protein 18..390
FT /label= Toso
FT Region 254..272
FT /label= transmembrane_region
XX WO9925832-A1.
XX 27-MAY-1999.
XX 16-NOV-1998; 98WO-US24391.
XX 17-AUG-1998; 98US-0135238.
XX 17-NOV-1997; 97US-0066063.
XX (STPD) UNIV LELAND STANFORD JUNIOR.
XX Hitoshi Y, Nolan GP;
XX WPI; 1999-338007/28.
XX N-PSDB; AAX76123.
XX DNA encoding Toso, a protein having inhibitory effects on TNF mediated apoptosis
XX Claim 20; Fig 2a; 70pp; English.
XX The present sequence is a Toso protein (I). (I) has anti-apoptotic and cytostatic activity. Toso (named after a Japanese liquor that is drunk on New Year's Day to celebrate long life and eternal youth) most likely acts by induction of cFLIP expression which inhibits caspase-8 processing. Recombinant (I) can be used to modulate apoptosis in a cell or to treat an apoptosis related condition in a mammal. Apoptosis related conditions can also be treated by administration of the Toso protein or antibody. Apoptosis related or mediated conditions that can be treated include diseases characterized by T cell overactivity, e.g. Sjogrens connective tissue disorder, autoimmune diseases, diseases where T cells actively destroy cells, including transplant rejection and conditions where cells of any kind that are not dying express Toso appropriately, e.g. cancer of T or B cell origin (where increased apoptosis would be appropriate).

XX SQ Sequence 390 AA;
Query Match 99.6%; Score 2047; DB 20; Length 390;
Best Local Similarity 99.7%; Pred. No. 1.8e-166;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MDRWLWPLYFLPVSGALRLPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCT 60
|||
Db 1 MDFWLWPLYFLPVSGALRLPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCT 60
Qy 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTL 120
|||
Db 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTL 120
Qy 121 NVHSEYEPSWEQPMPTPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVHHSSP 180
|||
Db 121 NVHSEYEPSWEQPMPTPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVHHSSP 180
Qy 181 TQITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQTPSYNHHTRLHRQALD 240
|||
Db 181 TQITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQTPSYNHHTRLHRQALD 240
Qy 241 YGSQSGREGQGHILPTILGLFLALLGLVVKRAVERRKALSRARRLAVRRALESSQ 300
|||
Db 241 YGSQSGREGQGHILPTILGLFLALLGLVVKRAVERRKALSRARRLAVRRALESSQ 300
Qy 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360
|||
Db 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360

QY 361 KTSCEYVSLYHQPAAMMEDSDSDYINVPA 390
ID AAY05001
Db 361 KTSCEYVSLYHQPAAMMEDSDSDYINVPA 390

RESULT 3
AA05001
ID AAY05001 standard; Protein; 390 AA.

XX AAY05001;
XX 16-JUN-1999 (first entry)
DT Human PIGRL-1 protein sequence.

DE PIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;
KW X-linked Severe Combined Immunodeficiency; XSCID; IgA deficiency;
KW diagnosis; therapy.

XX Homo sapiens.
XX EP905238-A2.
XX 31-MAR-1999.
XX 14-AUG-1998; 98EP-0306487.
XX 30-OCT-1997; 97US-0961564.
XX 25-AUG-1997; 97US-0056935.

XX (SMK) SMITHKLINE BEECHAM CORP.
XX Sweet RW, Truneh A, Wu S;
XX WPI; 1999-192666/17.
XX N-PSDB; AAX281178.
XX New polypeptides encoding PIGRL-1 useful for treating diseases such
XX as X-linked Severe Combined Immunodeficiency
XX Claim 11; Page 7; 26pp; English.

CC This sequence is the human PIGRL-1 protein of the invention.
CC Autoimmune diseases involving altered expression or activity of PIGRL-1
CC may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined
CC Immunodeficiency (XSCID) and IgA deficiency. These diseases can be
CC diagnosed or susceptibility to them predicted by: (1) determining whether
CC there is a mutation in the genomic copy of the gene encoding PIGRL-1; or
CC (2) measuring the amount of PIGRL-1 in a sample derived from the patient.
CC Patients deficient in PIGRL-1 can be treated by administering either the
CC PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient.
CC Patients with excessive expression or activity of PIGRL-1 can be treated
CC by administering an antagonist of PIGRL-1, an antisense nucleic acid
CC molecule which inhibits the expression of PIGRL-1 or administering
CC sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can
CC be used to identify its agonists by contacting a cell expressing PIGRL-1
CC with a candidate compound in the presence of a signal system and noting
CC the candidate as an agonist if a signal is produced. The same method can
CC be used to identify antagonists of PIGRL-1 but the presence of an
CC antagonist is indicated by a decrease in production of the signal.
CC Antibodies against PIGRL-1 may be used to isolate or identify clones
CC expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to
CC identify chromosomal mutations in the gene encoding PIGRL-1 in patients.
CC This information may then be correlated with the incidence of autoimmune
CC disease in those patients to identify whether the mutation causes the
CC disease.

XX Sequence 390 AA;

XX Query Match 99.6%; Score 2047; DB 20; Length 390;
XX Best Local Similarity 99.7%; Pred. No. 1.8e-166;
XX Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLWPLFLPVSGALRILPEVKVEGELGGSVTIKCPLPEMHVRIYLCREMAGSGTCGT 60
Db 1 MDRWLWPLFLPVSGALRILPEVKVEGELGGSVTIKCPLPEMHVRIYLCREMAGSGTCGT 60
QY 61 VVSTTNFKAKEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTL 120
Db 61 VVSTTNFKAKEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTL 120
QY 121 NVHSEYEPSEWQPMPEPKWFHLPYLFQMPAYASSSKFVTRVTPAQRKGVPVHHSSP 180
Db 121 NVHSEYEPSEWQPMPEPKWFHLPYLFQMPAYASSSKFVTRVTPAQRKGVPVHHSSP 180
QY 181 TTQITHRPRVSRASSVAGDKPTELPTSTASKISALEGLLKPTQPSYNHHTLHQRALD 240
Db 181 TTQITHRPRVSRASSVAGDKPTELPTSTASKISALEGLLKPTQPSYNHHTLHQRALD 240
QY 241 YGSQSGREGQGFHILPTITLGLFLLALLGLVYKRAVERRKALSRARRLAVEMRALESSQ 300
Db 241 YGSQSGREGQGFHILPTITLGLFLLALLGLVYKRAVERRKALSRARRLAVEMRALESSQ 300
QY 301 RPRGSPRPRSQNNIYSACPRRARGADAAGTGEAPVPGCAPLPAPLOVSESPWLHAPSL 360
Db 301 RPRGSPRPRSQNNIYSACPRRARGADAAGTGEAPVPGCAPLPAPLOVSESPWLHAPSL 360
QY 361 KTSCEYVSLYHQPAAMMEDSDSDYINVPA 390
Db 361 KTSCEYVSLYHQPAAMMEDSDSDYINVPA 390

RESULT 4
AAE05349
ID AAE05349 standard; Protein; 422 AA.

XX AAE05349;

DT 12-SEP-2001 (first entry)

DE Mouse Toso protein.

KW Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;
KW wound healing; immune response; vaccine; cancer; asthma; allergy;
KW cell trafficking; therapy; secreted protein; Fas-induced apoptosis;
KW Toso.

XX Mus sp.

XX WO200148192-A1.

XX 05-JUL-2001.

XX 21-DEC-2000; 2000WO-NZ00256.

XX 23-DEC-1999; 99US-0171678.

XX 28-NOV-2000; 2000US-0724864.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Murison JG;

XX WPI; 2001-425665/45.

XX N-PSDB; AAD10117.

XX Novel isolated polypeptide useful to isolate corresponding interacting
XX proteins or other compounds, to quantitatively determine levels of
XX interacting proteins or other compounds, and as therapeutic target -

XX Claim 6; Page 78-79; 101pp; English.

XX The patent discloses novel polynucleotides and their corresponding
XX proteins which play a major role in induction of growth, cell migration
XX and proliferation, cell-cell interaction and the differentiation of
XX tissue-specific cells. These proteins are important in the maintenance
XX of tissue integrity and thus are important in wound healing. They are

CC useful in various assays to determine the biological activity, to raise
CC antibodies, to isolate corresponding interacting proteins or other
CC compounds, to quantitatively determine levels of interacting proteins or
CC other compounds, and as therapeutic target in a whole range of disease
CC states. Compositions comprising the novel proteins of the invention are
CC useful for treating mammalian disorders. Polynucleotides of the invention
CC are useful in genome and physical mapping, in positional cloning of
CC genes, to tag or identify an organism or its reproductive material (as
CC non-disruptive tags for marking organisms), and for the diagnosis and
CC treatment of mammalian diseases which is the consequence of inappropriate
CC expression of kinase genes. They are useful for promoting immune response
CC as part of a vaccine or anti-cancer treatment, as target for cancer
CC treatment, as immunoregulatory and anti-inflammatory molecule, as
CC diagnostic for specific types of cancer and for development of an
CC anti-cancer treatment, and as a target for antagonists in the treatment
CC of diseases such as asthma and allergy. They are also useful to inhibit
CC or enhance the activity of the soluble molecule that binds proteins of
CC the invention, for tissue and neural regeneration, to promote or block
CC cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.
CC The present sequence is Toso, a secreted protein from mouse. Toso is
CC a cell surface, specific regulator of Fas-induced apoptosis in T-cells.
XX
XX

XX Sequence 422 AA;

Query Match 56.3%; Score 1157; DB 22; Length 422;
Best Local Similarity 57.5%; Pred. No. 2e-90;
Matches 234; Conservative 46; Mismatches 107; Indels 20; Gaps 6;

QY 1 MDRWLWLPFLPVSGALRLPEVKVGEGLGSGVTIKCPLEPMHVRIYLCREMAGSGTGT 60
DB 1 MDFWMLLLYFLPVSGALRLPEVQLNVEWEGSIIIECPQLHVRMYLCRQAKPGICST 60
QY 61 VYSTTNFKAEGYKGRVTLKQYPRKNFLPEVLTQTESDGYACGAGMNTDRGKTQVTL 120
DB 61 VYSNT-FVKXEYERRVTLPCLDKLLFVEMTQLTENDGGIYACGVGMKTDGKTQKITL 119
QY 121 NVHSEY-EPSEWEQPMPTPKFWFLPYLFQMP-----AYASSSFVTRVTTAQRGKVP 174
DB 120 NVHNEYPEPWEDEWTSRPRWLHRELFQHPMLHSGSEHPSGVIKAVTTPASKTEAPP 179
QY 175 VHSSTPTQITHRPVRSSVAGDKPRTFLPSTTASKTSALEGLLKPPQPSVNNHTRLH 234
DB 180 VHQPSITSVTQPRVIRAFVSATKSPALLPATPATSTQQA-IRPLEASYSHTRLH 238
QY 235 QRALDYGSGSGREGGQ-----FHILIPTILGLFLLALLGLVVKRERRKALSRAR 287
DB 239 EQTRRHGPHYGREDRGLHIPIEFHILPTFLGLLILLVLLGLVVKRAIQRRASSRRAG 298
QY 288 RLAVRWRALESS-----QPRGSPRPRSQNNIYSACPRARGADAAGTGPVPGPGAPL 342
DB 299 RLAMRRRGASRPFTQRRDAPQPRSQNNIYSACPRARGPDSLGPAPLLNAPASA 358
QY 343 PPAPQVSESPWLHAPSLKTSCEYVSLYHQPAAAMDDSDDDYINVP 389
DB 359 SPASQVLEAPWHPHTSLKMSCEYVSLGYPAVNLEDPSDDYINIP 405

RESULT 5
AAY05002
ID AAY05002 standard; peptide; 107 AA.
XX
XX AAY05002;
XX
XX 16-JUN-1999 (first entry)
XX
XX Human PIGRL-1 protein sequence fragment.

XX PIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;
KW X-linked Severe Combined Immunodeficiency; XSCID; IgA deficiency;
KW diagnosis; therapy.
XX
XX Homo sapiens.
OS
XX

PN EP905238-A2.
XX 31-MAR-1999.
XX 14-AUG-1998; 98EP-0306487.
XX 30-OCT-1997; 97US-0061564.
PR 25-AUG-1997; 97US-0056935.
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA Sweet RW, Truneh A, Wu S;
PI WPI; 1999-192666/17.
XX N-PSDB; AAX28179.
DR
XX New polypeptides encoding PIGRL-1 useful for treating diseases such
PT as X-linked Severe Combined Immunodeficiency
XX
PS Disclosure; Page 9; 26pp; English.
XX
XX This sequence is a fragment of the human PIGRL-1 of the invention.
CC Autoimmune diseases involving altered expression or activity of PIGRL-1
CC may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined
CC Immunodeficiency (XSCID) and IgA deficiency. These diseases can be
CC diagnosed or susceptibility to them predicted by: (1) determining whether
CC there is a mutation in the genomic copy of the gene encoding PIGRL-1;
CC (2) measuring the amount of PIGRL-1 in a sample derived from the patient.
CC Patients deficient in PIGRL-1 can be treated by administering either the
CC PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient.
CC Patients with excessive expression or activity of PIGRL-1 can be treated
CC by administering an antagonist of PIGRL-1, an antisense nucleic acid
CC molecule which inhibits the expression of PIGRL-1 or administering
CC sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can
CC be used to identify its agonists by contacting a cell expressing PIGRL-1
CC with a candidate compound in the presence of a signal system and noting
CC the candidate as an agonist if a signal is produced. The same method can
CC be used to identify antagonists of PIGRL-1 but the presence of an
CC antagonist is indicated by a decrease in production of the signal.
CC Antibodies against PIGRL-1 may be used to isolate or identify clones
CC expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to
CC identify chromosomal mutations in the gene encoding PIGRL-1 in patients.
CC This information may then be correlated with the incidence of autoimmune
CC disease in those patients to identify whether the mutation causes the
CC disease.
XX
XX Sequence 107 AA;

Query Match 26.8%; Score 550; DB 20; Length 107;
Best Local Similarity 99.0%; Pred. No. 2.2e-39;
Matches 104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLWLPFLPVSGALRLPEVKVGEGLGSGVTIKCPLEPMHVRIYLCREMAGSGTGT 60
DB 1 MDFWMLPLFLPVSGALRLPEVKVGEGLGSGVTIKCPLEPMHVRIYLCREMAGSGTGT 60
QY 61 VYSTTNFKAEGYKGRVTLKQYPRKNFLPEVLTQTESDGYACG 105
DB 61 VYSTTNFKAEGYKGRVTLKQYPRKNFLPEVLTQTESDGYACG 105

RESULT 6
AAY65401
ID AAY65401 standard; Protein; 97 AA.
XX
XX AAY65401;
XX
XX 01-FEB-2000 (first entry)
XX
XX Human 5' EST related polypeptide SEQ ID NO:1562.

XX Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
XX

QY 10 FLVPVSGALRILPEVKVEGELGSGVTIKCPLPE-----MHVRIYLCREMAGSGTCGTWVSTT 65
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 13 FSGVSTKSPIFGPOEVSIEGDSVITCYYPDTSVNRHTRKYWCROGA-SGMCTTLLSSN 71
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 66 NFIAEYKGRVILKQYPRKNLFLVEVTQLTESDSGVYACAGMNTDRGKTOKVTLNVHSE 125
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 72 GYLSKEYSGRANLINFENNTFVINEIQTQDDTGSYKCGLG-TSNRGLSFDVSLEV--- 127
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 126 YEPSWEQPMPTPKWFHLYLFQMPAYASSSKFVTRVTPAQRGKVP 173
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 128 -----SQVPELPSDTHV-----YTKDIGRNVITIECPFKRENVP 160
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
RESULT 8
AAG65698
ID AAG65698 standard; protein; 771 AA.
XX
AC AAG65698;
XX
DT 07-JAN-2002 (first entry)
XX
XX Mouse polymeric immunoglobulin receptor (pIgR) sequence.
XX
DE Polymetric immunoglobulin receptor; pIgR; ligand; therapeutic;
XX
KW carcinoma diagnosis; veterinary; mouse.
XX
OS Mus sp.
XX
XX WO200172846-A2.
XX
PD 04-OCT-2001.
XX
XX 26-MAR-2001; 2001WO-US09699.
XX
XX 27-MAR-2000; 2000US-192197P.
PR
PR 27-MAR-2000; 2000US-192198P.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Mostov KE, Chapin SJ, Richman-Eisenstat J;
PI
XX WPI; 2001-611619/70.
XX
XX New ligands binding to a specific region of a polymeric immunoglobulin
PT receptor, useful for transporting therapeutic or diagnostic
PT compositions into or across cells expressing pIgR e.g. in drug delivery
PT
XX
PS Disclosure; Fig 1; 102pp; English.
XX
XX The invention provides ligands that bind specifically to a region of an
CC animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves
CC to produce a stalk region remaining attached to the cell and a secretory
CC component existing in the organ of interest in several forms. The ligands
CC do not bind to the stalk or the most abundant form of the secretory
CC component present in the organ under physiological conditions. The
CC ligands are useful for transporting therapeutic or diagnostic
CC compositions into or across cells expressing pIgR, useful to introduce
CC or transport ligands such as antibodies and/or to deliver biologically
CC active components such as proteins, nucleic acids or detectable labels.
CC They are used to deliver therapeutic compositions to mucosal surfaces
CC such as the gastro-intestinal tract, respiratory system etc. in humans.
CC They are also useful to label cells expressing pIgR, e.g. to distinguish
CC epithelial cells from a mixed cell population in pathology studies or to
CC aid in carcinoma diagnosis (since pIgR expression is reduced in
CC carcinomas relative to normal epithelium). They can also be used to
CC deliver veterinary compositions, especially in mammals such as farm,
CC domestic or wild mammals or birds e.g. birds reared for human
CC consumption. The present sequence represents a mouse pIgR sequence.
XX
XX Sequence 771 AA;
SQ

Query Match 9.5%; Score 195; DB 22; Length 771;
Best Local Similarity 34.5%; Pred. No. 6,1e-08;
Matches 48; Conservative 21; Mismatches 54; Indels 16; Gaps 4;
QY 10 FLVPVSGALRILPEVKVEGELGSGVTIKCPLPE-----MHVRIYLCREMAGSGTCGTWVSTT 65
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 13 FSGVSTKSPIFGPOEVSIEGDSVITCYYPDTSVNRHTRKYWCROGA-SGMCTTLLSSN 71
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 66 NFIAEYKGRVILKQYPRKNLFLVEVTQLTESDSGVYACAGMNTDRGKTOKVTLNVHSE 125
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 72 GYLSKEYSGRANLINFENNTFVINEIQTQDDTGSYKCGLG-TSNRGLSFDVSLEV--- 127
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 126 YEPSWEQPMPTPKWFHLYLFQMPAYASSSKFVTRVTPAQRGKVP 144
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 128 -----SQVPELPSDTHV 139
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
RESULT 9
AAW03181
ID AAW03181 standard; Protein; 769 AA.
XX
AC AAW03181;
XX
DT 24-FEB-1997 (first entry)
XX
XX Rat poly-immunoglobulin receptor.
DE
XX
KW Rat; immunoglobulin; receptor; protection protein; mutants;
KW heavy chain; antigen binding domain; protection; pathogen;
KW mucosal; environment; gastrointestinal; passive; immunisation;
KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
KW poly; sorbinus.
XX
OS Rattus rattus.
XX
XX Key Location/Qualifiers
FH 13..45
PT Region /note= "putative immunoglobulin binding residues
FT Of domain I"
FT
FT Domain 1..120
FT /label= domain_I
FT Domain 110..230
FT /label= domain_II
FT Domain 210..340
FT /label= domain_III
FT Domain 320..450
FT /label= domain_IV
FT Domain 440..550
FT /label= domain_V
FT 550..606
FT /note= "external portions of domain VI"
FT Region 550..627
FT /note= "external portions of domain VI"
FT Region 625..660
FT /label= transmembrane_segment
FT Region 650..769
FT /label= intracellular_portion
XX
XX WO9621012-A1.
PN
XX 11-JUL-1996.
XX
XX 27-DEC-1995; 95WO-US16889.
XX
XX 04-MAY-1995; 95US-0434000.
PR
PR 30-DEC-1994; 94US-0367395.
XX
XX (PLAN-) PLANT BIOTECHNOLOGY INC.
PA (UNMF-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.
XX (PLAN-) PLANET BIOTECHNOLOGY INC.
XX
XX Hiatt AC, Lehner T, Ma JKC;
PI
XX

DR WPI; 1996-333987/33.
 DR N-PSDB; AAT31291.
 XX Immunoglobulin and protection protein complex and its prodn. in
 PT plants - useful for passive immunisation against mucosal antigens,
 PT esp. against S. mutans and S. sorbinus to prevent dental caries
 XX
 PS Disclosure; Pages 123-127; 152pp; English.
 XX
 CC The present sequence is the rat poly-immunoglobulin (Ig)
 CC receptor, a portion of which corresp. to residues 1-627, pref.
 CC 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,
 CC 440-550, 550-606 or 550-627 comprises a protection protein (PP).
 CC The Ig of the invention comprises a PP as above in association with
 CC an Ig derived heavy chain, having at least a portion of an antigen
 CC (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.
 CC gastrointestinal, environments, therefore enhancing its
 CC effectiveness in passively immunising animals against mucosal
 CC pathogens. The Ag binding domain is specifically derived from the
 CC Guy's 13 antibody, and the Ig can be used to prevent dental caries
 CC by binding, e.g. Streptococcus mutans serotypes c, e and f, or
 CC S. sorbinus serotypes d and g.
 XX
 XX Sequence 769 AA;
 SQ
 Query Match 9.0%; Score 185; DB 17; Length 769;
 Best Local Similarity 34.5%; Pred. No. 4.4e-07;
 Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
 QY 30 GGSVTIKCPLE----MHVRIYLCREMAGSGTCGTVVSTNFKAEYKGRVTLKQYPRKN 85
 DB 33 GNSVSITCYYPDTSVNRHTRKYWCROGA-NGYCATLISSNGYLSKEYSGRASLINFPE 91
 QY 86 LFLVEVTQLTESDGSVYACGAGMNTDRGKTQKVTNLNVHSEYPSWEEQPMPTPKWFHL 144
 DB 92 TFINIAHLTQEDTGSYKCGLG-TTNRGLFFDYSLEV-----SQVPEPNDTHV 139

RESULT 10
 AAG5697
 ID AAG65697 standard; protein; 769 AA.
 AC AAG65697;
 XX
 XX 07-JAN-2002 (first entry)
 DE Rat polymeric immunoglobulin receptor (pIgR) sequence.
 XX
 XX Polymeric immunoglobulin receptor; pIgR; ligand; therapeutic;
 KW carcinoma diagnosis; veterinary; rat.
 XX
 XX Rattus sp.
 XX
 XX WO200172846-A2.
 FN
 XX
 XX 04-OCT-2001.
 PD
 XX
 XX 26-MAR-2001; 2001WO-US09699.
 PE
 XX
 XX 27-MAR-2000; 2000US-192197P.
 PR
 XX 27-MAR-2000; 2000US-192198P.
 PR
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX
 XX Mostov KE, Chapin SJ, Richman-Eisenstat J;
 PI
 XX
 XX WPI; 2001-611619/70.
 DR
 XX
 XX New ligands binding to a specific region of a polymeric immunoglobulin
 PT receptor, useful for transporting therapeutic or diagnostic
 PT compositions into or across cells expressing pIgR e.g. in drug delivery
 PT
 XX

PS Disclosure; Fig 1; 102pp; English.
 XX
 CC The invention provides ligands that bind specifically to a region of an
 CC animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves
 CC to produce a stalk region remaining attached to the cell and a secretory
 CC component existing in the organ of interest in several forms. The ligands
 CC do not bind to the stalk or the most abundant form of the secretory
 CC component present in the organ under physiological conditions. The
 CC ligands are useful for transporting therapeutic or diagnostic
 CC compositions into or across cells expressing pIgR, useful to introduce
 CC or transport ligands such as antibodies, nucleic acids or detectable labels.
 CC active components such as proteins, nucleic acids or detectable labels.
 CC They are used to deliver therapeutic compositions to mucosal surfaces.
 CC such as the gastro-intestinal tract, respiratory system etc. in humans.
 CC They are also useful to label cells expressing pIgR, e.g. to distinguish
 CC epithelial cells from a mixed cell population in pathology studies or to
 CC aid in carcinoma diagnosis (since pIgR expression is reduced in
 CC carcinomas relative to normal epithelium). They can also be used to
 CC deliver veterinary compositions, especially in mammals such as farm,
 CC domestic or wild mammals or birds e.g. birds reared for human
 CC consumption. The present sequence represents a rat pIgR sequence.
 XX
 XX Sequence 769 AA;
 SQ
 Query Match 9.0%; Score 185; DB 22; Length 769;
 Best Local Similarity 34.5%; Pred. No. 4.4e-07;
 Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
 QY 30 GGSVTIKCPLE----MHVRIYLCREMAGSGTCGTVVSTNFKAEYKGRVTLKQYPRKN 85
 DB 33 GNSVSITCYYPDTSVNRHTRKYWCROGA-NGYCATLISSNGYLSKEYSGRASLINFPE 91
 QY 86 LFLVEVTQLTESDGSVYACGAGMNTDRGKTQKVTNLNVHSEYPSWEEQPMPTPKWFHL 144
 DB 92 TFINIAHLTQEDTGSYKCGLG-TTNRGLFFDYSLEV-----SQVPEPNDTHV 139

RESULT 11
 AAW50033
 ID AAW50033 standard; Protein; 532 AA.
 XX
 XX AAW50033;
 AC
 XX
 XX 26-JUN-1998 (first entry)
 DT
 XX
 XX Human immunity related factor.
 DE
 XX
 XX Lymph node; human; immunity related factor; research; treatment;
 KW immune disease; infectious disease.
 KW
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..16
 FT /label= sig_peptide
 FT Peptide 17..532
 FT /label= mat_peptide
 FT
 XX JP10072495-A.
 PN
 XX
 XX 17-MAR-1998.
 PD
 XX
 XX 11-JUN-1997; 97JP-0153218.
 PF
 XX
 XX 13-JUN-1996; 96JP-0152362.
 PR
 XX
 XX (ASAH) ASahi KASEI KOGYO KK.
 PA
 XX
 XX WPI; 1998-234766/21.
 DR
 XX N-PSDB; AAV20383.
 DR
 XX Immunity related factor - useful in the treatment of immune related
 PT and infectious diseases
 PT

CC immunoglobulin A (sIgA). It can be used as part of a method for
CC the production of sig molecules. This method is useful for
CC producing commercial quantities of sig (especially sIgA) to treat
CC or prevent infections. In particular, sIgA produced by the method
CC can be used to prevent or treat infections in mammals, birds or
CC fish; especially systemic infections or infections at a mucosal
CC surface. It is especially useful to prevent or treat infection
CC with human immunodeficiency virus (HIV), respiratory syncytial
CC virus, flu virus or cold virus. The method allows production of
CC commercial quantities of sig molecules for therapeutic use, not
CC previously possible; production using non-plant cells and a
CC single cell type is more efficient than a previous multi-step
CC process of fusing recombinant plant cells, and avoids alterations
CC of the sig by plant cells. sIgA molecules are more stable
CC and resistant to proteolysis than previously used IgA molecules,
CC and can be administered to prevent as well as to treat infections,
CC unlike e.g. IgG and IgM molecules.

XX Sequence 607 AA;

Query Match 8.9%; Score 182; DB 20; Length 607;
Best Local Similarity 37.4%; Pred. No. 5.7e-07;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PVSGALRLPEVKVEGELGSGVTIKCLPLP-----EMHVRVYLCREMAGSGTGTWVSTNPF 67
DB 21 PIFG-----PE-EVNSVEGNSVITCYPPTSVNRHTRKYWCROGARGG-CITLISSEGY 73
QY 68 IKAERYKGRVTLKOYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTLVN 122
DB 74 VSSKYAGRANLTNFPENGTFFVNNIAQLSQDSDSGRYKCGLGINS-RGLSFDVSLV 127

RESULT 14

AAW03178
ID AAW03178 standard; Protein; 746 AA.

XX AC AAW03178;

DT 24-FEB-1997 (first entry)

XX Human poly-immunoglobulin receptor.

XX Human; immunoglobulin; receptor; protection protein; mutants;
KW heavy chain; antigen binding domain; protection; pathogen;
KW mucosal; environment; gastrointestinal; passive; immunisation;
KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
XX poly; sorbinus.

OS Homo sapiens.

XX Key Location/Qualifiers
FH Region 13..45
FT /note= "putative immunoglobulin binding residues
of domain I"

FT Domain 1..120

FT /label= domain_I

FT Domain 110..230

FT /label= domain_II

FT Domain 210..340

FT /label= domain_III

FT Domain 320..450

FT /label= domain_IV

FT Domain 440..550

FT /label= domain_V

FT Region 550..606

FT /note= "external portions of domain VI"

FT Region 550..627

FT /note= "external portions of domain VI"

FT Region 625..660

FT /label= transmembrane_segment

FT Region 650..746

FT /label= intracellular_portion

XX WO9621012-A1.
PN 11-JUL-1996.
XX 27-DEC-1995; 95WO-US16889.
XX 04-MAY-1995; 95US-0434000.
PR 30-DEC-1994; 94US-0367395.
XX (PLAN-) PLANT BIOTECHNOLOGY INC.
PA (UNME-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.
PA (PLAN-) PLANET BIOTECHNOLOGY INC.
XX Hiatt AC, Lehner T, Ma JKC;
XX WPI; 1996-333987/33.
DR N-PSDB; AAT31288.
XX Immunoglobulin and protection protein complex and its prodn. in
PT plants - useful for passive immunisation against mucosal antigens,
PT esp. against S. mutans and S. sorbinus to prevent dental caries
XX Disclosure; Pages 105-108; 152pp; English.

XX The present sequence is the human poly-immunoglobulin (Ig)
CC receptor, a portion of which corresp. to residues 1-627, pref.
CC 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,
CC 440-550, 550-606 or 550-627 comprises a protection protein (PP).
CC The Ig of the invention comprises a PP as above in association with
CC an Ig derived heavy chain, having at least a portion of an antigen
CC (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.
CC gastrointestinal, environments, therefore enhancing its
CC effectiveness in passively immunising animals against mucosal
CC pathogens. The Ag binding domain is specifically derived from the
CC Guy's 13 antibody, and the Ig can be used to prevent dental caries
CC by binding, e.g. Streptococcus mutans serotypes c, e and f, or
CC S. sorbinus serotypes d and g.

XX Sequence 746 AA;

Query Match 8.9%; Score 182; DB 17; Length 746;
Best Local Similarity 37.4%; Pred. No. 7.5e-07;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PVSGALRLPEVKVEGELGSGVTIKCLPLP-----EMHVRVYLCREMAGSGTGTWVSTNPF 67
DB 3 PIFG-----PE-EVNSVEGNSVITCYPPTSVNRHTRKYWCROGARGG-CITLISSEGY 55

QY 68 IKAERYKGRVTLKOYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTLVN 122

DB 56 VSSKYAGRANLTNFPENGTFFVNNIAQLSQDSDSGRYKCGLGINS-RGLSFDVSLV 109

RESULT 15

AAW03179
ID AAW03179 standard; Protein; 757 AA.

XX AC AAW03179;

DT 24-FEB-1997 (first entry)

XX Bovine poly-immunoglobulin receptor.

DE Bovine; immunoglobulin; receptor; protection protein; mutants;
KW heavy chain; antigen binding domain; protection; pathogen;
KW mucosal; environment; gastrointestinal; passive; immunisation;
KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
XX poly; sorbinus; cow.

OS Bos taurus.

XX Key Location/Qualifiers

```

FT Region 13..45
FT /note= "putative immunoglobulin binding residues
FT of domain I"
FT
FT Domain 1..120
FT /label= domain_I
FT Domain 110..230
FT /label= domain_II
FT Domain 210..340
FT /label= domain_III
FT Domain 320..450
FT /label= domain_IV
FT Domain 440..550
FT /label= domain_V
FT Region 550..606
FT /note= "external portions of domain VI"
FT Region 550..627
FT /note= "external portions of domain VI"
FT Region 625..660
FT /label= transmembrane_segment
FT Region 650..757
FT /label= intracellular_portion
FT
XX
PN W09621012-A1.
XX
XX 11-JUL-1996.
XX
XX 27-DEC-1995; 95WO-US16889.
XX
XX 04-MAY-1995; 95US-0434000.
XX 30-DEC-1994; 94US-0367395.
XX
XX (PLAN-) PLANT BIOTECHNOLOGY INC.
XX (UNNE-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.
XX (PLAN-) PLANET BIOTECHNOLOGY INC.
XX
XX Hiatt AC, Lehner T, Ma JKC;
XX
XX WPI: 1996-333987/33.
XX DR N-PSDB; AAT31289.
XX
XX Immunoglobulin and protection protein complex and its prodn. in
PT plants - useful for passive immunisation against mucosal antigens,
PT esp. against S. mutans and S. sorbinus to prevent dental caries
XX
XX Disclosure: Pages 111-115; 152pp; English.
XX
XX The present sequence is the bovine poly-immunoglobulin (Ig)
XX receptor, a portion of which corresp. to residues 1-627, pref.
XX 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,
XX 440-550, 550-606 or 550-627 comprises a protection protein (PP).
XX The Ig of the invention comprises a PP as above in association with
XX an Ig derived heavy chain, having at least a portion of an antigen
XX (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.
XX gastrointestinal, environments, therefore enhancing its
XX effectiveness in passively immunising animals against mucosal
XX pathogens. The Ag binding domain is specifically derived from the
XX Guy's 13 antibody, and the Ig can be used to prevent dental caries
XX by binding, e.g. Streptococcus mutans serotypes c, e and f, or
XX S. sorbinus serotypes d and g.
XX
XX Sequence 757 AA;
XX
Query Match 8.9%; Score 182; DB 17; Length 757;
Best Local Similarity 34.1%; Pred. No. 7.7e-07;
Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;
QY 17 LRILPVKVEGL-----GGSVTKICPLP-----EMHVIYLCREMAGSGTCTGVV 62
DB 10 LAIFPVVSMKSIIFGPEEVSVEGRSVKICYPTSVNHHTRKYWCROGA-QGRCTTII 68
QY 63 STNFIKAQYKGYTLKQYPRKNLFLVEVTQTESDGVYACGAGMNTDRGKTKQVTLNV 122
DB 69 SSGEYVSDDYVGRANLTNFPESGTFVVDISHLTHKDSGRYKGLGISS-RGLNFDVSLEV 127

```

QY 123 HSE 125
Db 128 SQD 130

Search completed: October 28, 2002, 17:30:51
Job time : 44.8871 secs